(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 April 2002 (04.04.2002)

PCT

(10) International Publication Number
WO 02/26822 A2

(51) International Patent Classification[7]: C07K 14/705

(21) International Application Number: PCT/US01/30290

(22) International Filing Date:
26 September 2001 (26.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/235,451 26 September 2000 (26.09.2000) US

(71) Applicant *(for all designated States except US)*: **GENENTECH, INC.** [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DEVAUX, Brigitte** [US/US]; 1113 Oliver Street, Redwood City, CA 94061 (US). **EBERHARD, David** [US/US]; #212, 360 Guerrero Street, San Francisco, CA 94103 (US). **GODDARD, Audrey** [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). **GODOWSKI, Paul, J.** [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). **GRIMALDI, J., Christopher** [US/US]; 1434 36th Avenue, San Francisco, CA 94122 (US). **HILLAN, Kenneth, J.** [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). **WATANABE, Colin, K.** [US/US]; 128 Corliss Drive, Moraga, CA 94556 (US). **WOOD, William, I.** [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US). **YANSURA, Daniel, G.** [US/US]; 330 Carmel Avenue, Pacifica, CA 94044 (US). **ZHANG, Zemin** [CN/US]; 876 Taurus Drive, Foster City, CA 94404 (US).

(74) Agents: **TAN, Lee, K.** et al.; Genentech, Inc., MS49, 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/26822 A2

(54) Title: PUMPCN COMPOSITIONS AND USES THEREOF

(57) Abstract: The present invention is directed to novel polypeptides Protein Upregulated in Metastatic Prostate Cancer (PUMPCn) and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

PUMPCn COMPOSITIONS AND USES THEREOF

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptide Protein Upregulated in Metastatic Prostate Cancer (PUMPCn), designated herein as "PRO23203" polypeptides.

BACKGROUND OF THE INVENTION

Membrane-bound proteins and receptors can play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesin molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

Efforts are being undertaken by both industry and academia to identify new, native receptor or membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins.

For men, prostate cancer is the second most fatal cancer after lung cancer. The increase in the incidence of prostate cancer is projected to be one of the highest worldwide. Indeed, 1 man in 6 will develop invasive prostate cancer. In advanced stages, prostate cancer

metastasizes to the bone, and prostate cancer is usually incurable once it has metastasized. Currently, prostate-specific antigen (PSA) is the most widely used tumor marker for screening, diagnosis, and monitoring prostate cancer. However, widespread use of PSA as a tool for screening is controversial since PSA fails to discriminate accurately between benign and malignant prostate disease.

Depending on the stage of the cancer, prostate and bladder cancer treatment involves one or a combination of the following therapies: surgery to remove cancerous tissue, radiation therapy, chemotherapy, androgen deprivation (e.g., hormonal therapy) in the case of prostate cancer. The majority of patients who undergo hormone therapy progress to develop androgen-independent disease. Currently, there is no effective treatment for the 20-40% of prostate cancer patients who develop recurrent disease after surgery or radiation therapy, or for those in whom the cancer has metastasized at the time of diagnosis. Chemotherapy has its toxic side effects, especially in elderly patients. Development of new forms of therapy especially for disease refractory to androgen deprivation is an urgent need in the management of prostatic carcinoma.

Antibody-based therapy has proved very effective in the treatment of various cancers. For example, HERCEPTIN® and RITUXAN® (both from Genentech, S. San Francisco), have been used successfully to treat breast cancer and non-Hodgkin's lymphoma, respectively.

The present invention provides alternative methods of treating cancer that overcome the limitations of conventional therapeutic methods as well as offer additional advantages that will be apparent from the detailed description below.

We herein describe the identification and characterization of novel polypeptides PUMPCn designated herein as PRO23203 polypeptides. Fragments of PUMPCn have been described, see WO 98/37093, WO 98/37418, WO 99/61469, WO 99/62941, WO 00/04149, WO 01/25272, U.S. 6,048,970, WO 99/62941 and WO 01/40276.

SUMMARY OF THE INVENTION

A cDNA clone (designated herein as DNA185171-2994) has been identified as PUMPCn and that encodes a novel polypeptide, designated in the present application as "PRO23203".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO23203 polypeptide.

In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence

having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO23203 polypeptide having the sequence of amino acid residues from about 1 to about 454, inclusive, of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a).

In another aspect, the isolated nucleic acid molecule comprises (a) a nucleotide sequence encoding a PRO23203 polypeptide having the sequence of amino acid residues from about 1 to about 454, inclusive, of Figure 2 (SEQ ID NO:2), or (b) the complement of the nucleotide sequence of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about

99% nucleic acid sequence identity to (a) a DNA molecule having the sequence of nucleotides from about 188 to about 1549, inclusive, of Figure 1 (SEQ ID NO:1), or (b) the complement of the DNA molecule of (a).

In another aspect, the isolated nucleic acid molecule comprises (a) the nucleotide sequence of from about 188 to about 1549, inclusive, of Figure 1 (SEQ ID NO:1), or (b) the complement of the nucleotide sequence of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by the human protein cDNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994) or (b) the complement of the DNA molecule of (a). In a preferred embodiment, the isolated nucleic acid molecule comprises (a) a nucleotide sequence encoding the same mature polypeptide encoded by the human protein cDNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994) or (b) the complement of the nucleotide sequence of (a).

In another aspect, the invention concerns an isolated nucleic acid molecule which comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity,

alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) the full-length polypeptide coding sequence of the human protein cDNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994) or (b) the complement of the nucleotide sequence of (a). In a preferred embodiment, the isolated nucleic acid molecule comprises (a) the full-length polypeptide coding sequence of the DNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994) or (b) the complement of the nucleotide sequence of (a).

In another aspect, the invention concerns an isolated nucleic acid molecule which encodes an active PRO23203 polypeptide as defined below comprising a nucleotide sequence that hybridizes to the complement of a nucleic acid sequence that encodes amino acids 1 to about 454, inclusive, of Figure 2 (SEQ ID NO:2). Preferably, hybridization occurs under stringent hybridization and wash conditions.

In yet another aspect, the invention concerns an isolated nucleic acid molecule which encodes an active PRO23203 polypeptide as defined below comprising a nucleotide sequence that hybridizes to the complement of the nucleic acid sequence between about nucleotides 188 and about 1549, inclusive, of Figure 1 (SEQ ID NO:1). Preferably, hybridization occurs under stringent hybridization and wash conditions.

In a further aspect, the invention concerns an isolated nucleic acid molecule having at least about 1204 nucleotides and which is produced by hybridizing a test DNA molecule under stringent conditions with (a) a DNA molecule encoding a PRO23203 polypeptide having the sequence of amino acid residues from about 1 to about 454, inclusive, of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a), and, if the test DNA molecule has at least about an 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity,

alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) or (b), and isolating the test DNA molecule.

Another aspect of the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO23203 polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain has been tentatively identified as extending from about amino acid position 210 to about amino acid position 230 in the sequence of Figure 2 (SEQ ID NO:2). Therefore, soluble extracellular domains of the herein described PRO23203 polypeptides are contemplated.

In this regard, another aspect of the present invention is directed to an isolated nucleic acid molecule which comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding amino acids 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA

molecule of (a). In a specific aspect, the isolated nucleic acid molecule comprises a nucleotide sequence which (a) encodes amino acids 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2), or (b) is the complement of the DNA molecule of (a).

Another embodiment is directed to fragments of a PRO23203 polypeptide coding sequence that may find use as, for example, hybridization probes or for encoding fragments of a PRO23203 polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO23203 antibody. Such nucleic acid fragments are usually at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 160 nucleotides in length, alternatively at least about 170 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 190 nucleotides in length, alternatively at least about 200 nucleotides in length, alternatively at least about 250 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 350 nucleotides in length, alternatively at least about 400 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 500 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 700 nucleotides in length, alternatively at least about 800 nucleotides in length, alternatively at least about 900 nucleotides in length and alternatively at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. In a preferred embodiment, the nucleotide sequence fragment is derived from any coding region of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1). It is noted that novel fragments of a PRO23203 polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO23203 polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO23203 polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO23203 polypeptide-encoding nucleotide sequences are

contemplated herein and can be determined without undue experimentation. Also contemplated are the PRO23203 polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO23203 polypeptide fragments that comprise a binding site for an anti-PRO23203 antibody.

In another embodiment, the invention provides a vector comprising a nucleotide sequence encoding PRO23203 or its variants. The vector may comprise any of the isolated nucleic acid molecules hereinabove identified.

A host cell comprising such a vector is also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing PRO23203 polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of PRO23203 and recovering PRO23203 from the cell culture.

In another embodiment, the invention provides isolated PRO23203 polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a specific aspect, the invention provides isolated native sequence PRO23203 polypeptide, which in certain embodiments, includes an amino acid sequence comprising residues from about 1 to about 454 of Figure 2 (SEQ ID NO:2).

In another aspect, the invention concerns an isolated PRO23203 polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to the sequence of amino acid residues from about 1 to about 454, inclusive, of Figure 2 (SEQ ID NO:2).

In a further aspect, the invention concerns an isolated PRO23203 polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity,

alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by the human protein cDNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994). In a preferred embodiment, the isolated PRO23203 polypeptide comprises an amino acid sequence encoded by the human protein cDNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994).

Another aspect the invention provides an isolated PRO23203 polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO23203 polypeptide and recovering the PRO23203 polypeptide from the cell culture.

As such, one aspect of the present invention is directed to an isolated soluble PRO23203 polypeptide which comprises an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92%

amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to amino acids 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2). In a preferred aspect, the isolated soluble PRO23203 polypeptide comprises amino acids 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2).

In yet another aspect, the invention concerns an isolated PRO23203 polypeptide, comprising the sequence of amino acid residues from about 1 to about 454, inclusive, of Figure 2 (SEQ ID NO:2), or a fragment thereof which is biologically active or sufficient to provide a binding site for an anti-PRO23203 antibody, wherein the identification of PRO23203 polypeptide fragments that possess biological activity or provide a binding site for an anti-PRO23203 antibody may be accomplished in a routine manner using techniques which are well known in the art. Preferably, the PRO23203 fragment retains a qualitative biological activity of a native PRO23203 polypeptide.

In a still further aspect, the invention provides a polypeptide produced by (i) hybridizing a test DNA molecule under stringent conditions with (a) a DNA molecule encoding a PRO23203 polypeptide having the sequence of amino acid residues from about 1 to about 454, inclusive, of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a), and if the test DNA molecule has at least about an 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity

and alternatively at least about 99% nucleic acid sequence identity to (a) or (b), (ii) culturing a host cell comprising the test DNA molecule under conditions suitable for expression of the polypeptide, and (iii) recovering the polypeptide from the cell culture.

In another embodiment, the invention provides chimeric molecules comprising a PRO23203 polypeptide fused to a heterologous polypeptide or amino acid sequence, wherein the PRO23203 polypeptide may comprise any PRO23203 polypeptide, variant or fragment thereof as hereinbefore described. An example of such a chimeric molecule comprises a PRO23203 polypeptide fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody as defined below which specifically binds to a PRO23203 polypeptide as hereinbefore described. Optionally, the antibody is a monoclonal antibody, an antibody fragment or a single chain antibody. In one aspect of the invention, the anti-PUMPCn antibody binds to an epitope present within amino acids 1-50 of SEQ ID NO. 2, alternatively the epitope is present within amino acids 51-100, or amino acids 101-150, or 151-200, or 201-250, or 251-300. The invention also provides a monoclonal antibody that binds an epitope present within amino acids 25-213 of PUMPCn [see SEQ ID NO. 2]. In a specific embodiment, the mAbs 3248 though 3256 described in Example 13 and having ATCC accession numbers __ are provided.

In one embodiment, the invention provides an antibody that binds to PUMPCn on a live cell. In another embodiment, the invention provides an antibody that binds to an extracellular domain of PUMPCn. The invention provides isolated anti-PUMPCn antibodies that internalize upon binding to PUMPCn on a mammalian cell in vivo. These antibodies can also target a PUMPCn-expressing tumor cell in vivo. In a specific embodiment, the anti-PUMPCn antibodies internalize upon binding to PUMPCn on cancer cells, including prostate cancer and lung cancer. Also provided are antibodies that compete for binding to the same epitope as the epitope bound by any of the monoclonal antibodies of the invention. In another embodiment, an isolated anti-PUMPCn monoclonal antibody that inhibits the growth of PUMPCn-expressing cancer cells *in vivo*, or is cytotoxic *in vivo*, to such cells and tumors containing such cells, is provided.

The invention also provides anti-PUMPCn antibodies that are conjugated to a cytotoxic agent or to a growth inhibitory agent. The antibodies are internalizing and/or growth inhibitory antibodies. The cytotoxic agent can be a toxin, antibiotic, radioactive isotope or nucleolytic enzyme. In a preferred embodiment, the toxin is a maytansinoid, more preferably the maytansinoid having the structure shown in Figure 6.

The anti-PUMPCn antibodies of the preceding embodiments include intact (full

length) antibodies as well as antibody fragments. The antibodies of the invention include those produced in mammalian or bacterial cells.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO23203 polypeptide as defined below. In a particular embodiment, the agonist or antagonist is an anti-PRO23203 antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO23203 polypeptide which comprise contacting the PRO23203 polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO23203 polypeptide. Preferably, the PRO23203 polypeptide is a native PRO23203 polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO23203 polypeptide, or an agonist or antagonist of a PRO23203 polypeptide as herein described, or an anti-PRO23203 antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO23203 polypeptide, or an agonist or antagonist thereof as herein described, or an anti-PRO23203 antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO23203 polypeptide, an agonist or antagonist thereof or an anti-PRO23203 antibody.

Yet a separate aspect of the invention is a method of killing a PUMPCn-expressing cancer cell, comprising contacting the cancer cell with an anti-PUMPCn antibody of any of the above embodiments, thereby killing the cancer cell. Another aspect is a method of alleviating or treating a PUMPCn-expressing cancer in a mammal, comprising administering a therapeutically effective amount of the anti-PUMPCn antibodies of the invention to the mammal. In preferred embodiments of the preceding two methods, the cancer is a prostate, bladder or lung cancer, more preferably prostate cancer and especially an androgen independent prostate cancer cell or a metastatic prostate cancer. In a preferred embodiment of these methods, the anti-PUMPCn antibody is a human or a humanized antibody. In another preferred embodiment, the antibody is conjugated to a cytotoxic agent such as a toxin or a radioactive isotope. Preferably, the toxin is calicheamicin or a maytansinoid such as "DM1" having the structure shown in Fig. 6. The method of alleviating the PUMPCn-expressing cancer anticipates administration of the anti-PUMPCn antibody in conjuction with chemotherapy wherein the mammal is also receiving at least one chemotherapeutic agent. In a specific embodiment, the chemotherapeutic agent is a taxane such as paclitaxel (TAXOL®) or docetaxel, or derivatives and analogs thereof.

In a further aspect, the invention provides an article of manufacture comprising a container and a composition contained therein, wherein the composition comprises an anti-PUMPCn antibody of the above embodiments, and further comprising a package insert indicating that the composition can be used to alleviate or treat a PUMPCn-expressing cancer and in particular, prostate cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the nucleotide sequence (SEQ ID NO:1) of a cDNA containing a nucleotide sequence (nucleotides 188-1549) encoding native sequence PRO23203, wherein the nucleotide sequence (SEQ ID NO:1) is a clone designated herein as "DNA185171-2994". Also presented in bold font and underlined are the positions of the respective start and stop codons.

Figure 2 shows the amino acid sequence (SEQ ID NO:2) of a native sequence PRO23203 polypeptide as derived from the coding sequence of SEQ ID NO:1.

Figure 3 shows a GEPIS result for DNA185171-2994 [described in Example 1]. The type of tissue is indicated along the side of the figure. The relative abundance is plotted on the Y axis and determines if there is significant expression. The X axis lists the condition of the hit from that tissue, tumor (TUM), other (OTH), non-tumor tissue (NON), metastatic (MET), inflammatory (INF).

Figure 4 shows expression of PUMPCn relative to normal prostate in several human cDNA libraries, as described in Example 11.

Figure 5 shows the different fold expression of PUMPCn gene between tumor and normal tissue in human cDNA libraries , as described in Example 11.

Figure 6 shows the structure of the maytansinoid designated "DM1"

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "PRO23203 polypeptide", "PRO23203 protein" and "PRO23203" when used herein encompass native sequence PRO23203 and PRO23203 polypeptide variants (which are further defined herein). The PRO23203 polypeptide may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant and/or synthetic methods.

A "native sequence PRO23203" comprises a polypeptide having the same amino acid sequence as a PRO23203 derived from nature. Such native sequence PRO23203 can be

isolated from nature or can be produced by recombinant and/or synthetic means. The term "native sequence PRO23203" specifically encompasses naturally-occurring truncated or secreted forms (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the PRO23203. In one embodiment of the invention, the native sequence PRO23203 is a mature or full-length native sequence PRO23203 comprising amino acids 1 to 454 of Figure 2 (SEQ ID NO:2). Also, while the PRO23203 polypeptide disclosed in Figure 2 (SEQ ID NO:2) is shown to begin with the methionine residue designated herein as amino acid position 1, it is conceivable and possible that another methionine residue located either upstream or downstream from amino acid position 1 in Figure 2 (SEQ ID NO:2) may be employed as the starting amino acid residue for the PRO23203 polypeptide.

The PRO23203 polypeptide "extracellular domain(s)" or "ECD" refers to a form of the PRO23203 polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO23203 polypeptide ECD will have less than about 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than about 0.5% of such domains. It will be understood that any transmembrane domain(s) identified for the PRO23203 polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified. As such, in one embodiment of the present invention, the extracellular domain of a PRO23203 polypeptide comprises amino acids 1 to X, wherein X is any amino acid from amino acid 205 to 214 of Figure 2 (SEQ ID NO:2).

"PRO23203 variant polypeptide" means an active PRO23203 polypeptide as defined below having at least about 80% amino acid sequence identity with the amino acid sequence of (a) residues 1 to 454 of the PRO23203 polypeptide shown in Figure 2 (SEQ ID NO:2), (b) 1 to X of Figure 2 (SEQ ID NO:2), wherein X is any amino acid from amino acid 205 to 214 of Figure 2 (SEQ ID NO:2) or (c) another specifically derived fragment of the amino acid sequence shown in Figure 2 (SEQ ID NO:2). Such PRO23203 variant polypeptides include, for instance, PRO23203 polypeptides wherein one or more amino acid residues are added, or deleted, at the N- and/or C-terminus, as well as within one or more internal domains, of the sequence of Figure 2 (SEQ ID NO:2). Ordinarily, a PRO23203 variant polypeptide will have at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid

sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity with (a) residues 1 to 454 of the PRO23203 polypeptide shown in Figure 2 (SEQ ID NO:2), (b) 1 to X of Figure 2 (SEQ ID NO:2), wherein X is any amino acid from amino acid 205 to 214 of Figure 2 (SEQ ID NO:2) or (c) another specifically derived fragment of the amino acid sequence shown in Figure 2 (SEQ ID NO:2). PRO23203 variant polypeptides do not encompass the native PRO23203 polypeptide sequence. Ordinarily, PRO23203 variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length, alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more..

"Percent (%) amino acid sequence identity" with respect to the PRO23203 polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a PRO23203 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared. For purposes herein, however, % amino acid

sequence identity values are obtained as described below by using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

Table 1

```
/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int     _day[26][26] = {
/*    A  B  C  D  E  F  G  H  I  J  K  L  M  N  O  P  Q  R  S  T  U  V  W  X  Y  Z */
/* A */  { 2, 0,-2, 0, 0,-4, 1,-1,-1, 0,-1,-2,-1, 0,_M, 1, 0,-2, 1, 1, 0, 0,-6, 0,-3, 0},
/* B */  { 0, 3,-4, 3, 2,-5, 0, 1,-2, 0, 0,-3,-2, 2,_M,-1, 1, 0, 0, 0, 0,-2,-5, 0,-3, 1},
/* C */  {-2,-4,15,-5,-5,-4,-3,-3,-2, 0,-5,-6,-5,-4,_M,-3,-5,-4, 0,-2, 0,-2,-8, 0, 0,-5},
/* D */  { 0, 3,-5, 4, 3,-6, 1, 1,-2, 0, 0,-4,-3, 2,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 2},
/* E */  { 0, 2,-5, 3, 4,-5, 0, 1,-2, 0, 0,-3,-2, 1,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 3},
/* F */  {-4,-5,-4,-6,-5, 9,-5,-2, 1, 0,-5, 2, 0,-4,_M,-5,-5,-4,-3,-3, 0,-1, 0, 0, 7,-5},
/* G */  { 1, 0,-3, 1, 0,-5, 5,-2,-3, 0,-2,-4,-3, 0,_M,-1,-1,-3, 1, 0, 0,-1,-7, 0,-5, 0},
/* H */  {-1, 1,-3, 1, 1,-2,-2, 6,-2, 0, 0,-2,-2, 2,_M, 0, 3, 2,-1,-1, 0,-2,-3, 0, 0, 2},
/* I */  {-1,-2,-2,-2,-2, 1,-3,-2, 5, 0,-2, 2, 2,-2,_M,-2,-2,-2,-1, 0, 0, 4,-5, 0,-1,-2},
/* J */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */  {-1, 0,-5, 0, 0,-5,-2, 0,-2, 0, 5,-3, 0, 1,_M,-1, 1, 3, 0, 0, 0,-2,-3, 0,-4, 0},
/* L */  {-2,-3,-6,-4,-3, 2,-4,-2, 2, 0,-3, 6, 4,-3,_M,-3,-2,-3,-3,-1, 0, 2,-2, 0,-1,-2},
/* M */  {-1,-2,-5,-3,-2, 0,-3,-2, 2, 0, 0, 4, 6,-2,_M,-2,-1, 0,-2,-1, 0, 2,-4, 0,-2,-1},
/* N */  { 0, 2,-4, 2, 1,-4, 0, 2,-2, 0, 1,-3,-2, 2,_M,-1, 1, 0, 1, 0, 0,-2,-4, 0,-2, 1},
/* O */  {_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M, 0,_M,_M,_M,_M,_M,_M,_M,_M,_M,_
            M,_M},
/* P */  { 1,-1,-3,-1,-1,-5,-1, 0,-2, 0,-1,-3,-2,-1,_M, 6, 0, 0, 1, 0, 0,-1,-6, 0,-5, 0},
/* Q */  { 0, 1,-5, 2, 2,-5,-1, 3,-2, 0, 1,-2,-1, 1,_M, 0, 4, 1,-1,-1, 0,-2,-5, 0,-4, 3},
/* R */  {-2, 0,-4,-1,-1,-4,-3, 2,-2, 0, 3,-3, 0, 0,_M, 0, 1, 6, 0,-1, 0,-2, 2, 0,-4, 0},
/* S */  { 1, 0, 0, 0, 0,-3, 1,-1,-1, 0, 0,-3,-2, 1,_M, 1,-1, 0, 2, 1, 0,-1,-2, 0,-3, 0},
/* T */  { 1, 0,-2, 0, 0,-3, 0,-1, 0, 0, 0,-1,-1, 0,_M, 0,-1,-1, 1, 3, 0, 0,-5, 0,-3, 0},
/* U */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */  { 0,-2,-2,-2,-2,-1,-1,-2, 4, 0,-2, 2, 2,-2,_M,-1,-2,-2,-1, 0, 0, 4,-6, 0,-2,-2},
/* W */  {-6,-5,-8,-7,-7, 0,-7,-3,-5, 0,-3,-2,-4,-4,_M,-6,-5, 2,-2,-5, 0,-6,17, 0, 0,-6},
/* X */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */  {-3,-3, 0,-4,-4, 7,-5, 0,-1, 0,-4,-1,-2,-2,_M,-5,-4,-4,-3,-3, 0,-2, 0, 0,10,-4},
/* Z */  { 0, 1,-5, 2, 3,-5, 0, 2,-2, 0, 0,-2,-1, 1,_M, 0, 3, 0, 0, 0, 0,-2,-6, 0,-4, 4}
};

/*
 */
#include <stdio.h>
#include <ctype.h>

#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
#define MX          4       /* save if there's at least MX-1 bases since last jmp */

#define DMAT        3       /* value of matching bases */
#define DMIS        0       /* penalty for mismatched bases */
#define DINS0       8       /* penalty for a gap */
#define DINS1       1       /* penalty per base */
#define PINS0       8       /* penalty for a gap */
#define PINS1       4       /* penalty per residue */

struct jmp {
```

Table 1-(cont')

```
        short           n[MAXJMP];   /* size of jmp (neg for dely) */
        unsigned short  x[MAXJMP];   /* base no. of jmp in seq x */
};                                   /* limits seq to 2^16 -1 */

struct diag {
        int             score;       /* score at last jmp */
        long            offset;      /* offset of prev block */
        short           ijmp;        /* current jmp index */
        struct jmp      jp;          /* list of jmps */
};

struct path {
        int     spc;            /* number of leading spaces */
        short   n[JMPS];        /* size of jmp (gap) */
        int     x[JMPS];        /* loc of jmp (last elem before gap) */
};

char            *ofile;                 /* output file name */
char            *namex[2];              /* seq names: getseqs() */
char            *prog;                  /* prog name for err msgs */
char            *seqx[2];               /* seqs: getseqs() */
int             dmax;                   /* best diag: nw() */
int             dmax0;                  /* final diag */
int             dna;                    /* set if dna: main() */
int             endgaps;                /* set if penalizing end gaps */
int             gapx, gapy;             /* total gaps in seqs */
int             len0, len1;             /* seq lens */
int             ngapx, ngapy;           /* total size of gaps */
int             smax;                   /* max score: nw() */
int             *xbm;                   /* bitmap for matching */
long            offset;                 /* current offset in jmp file */
struct  diag    *dx;                    /* holds diagonals */
struct  path    pp[2];                  /* holds path for seqs */

char            *calloc(), *malloc(), *index(), *strcpy();
char            *getseq(), *g_calloc();


/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 *   where file1 and file2 are two dna or two protein sequences.
 *   The sequences can be in upper- or lower-case an may contain ambiguity
 *   Any lines beginning with ';', '>' or '<' are ignored
 *   Max file length is 65535 (limited by unsigned short x in the jmp struct)
 *   A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 *   Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static  _dbval[26] = {
        1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};
```

Table 1 (cont')

```
static  _pbval[26] = {
        1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
        128, 256, 0xFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
        1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
        1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                                    main
        int     ac;
        char    *av[ ];
{
        prog = av[0];
        if (ac != 3) {
                fprintf(stderr,"usage: %s file1 file2\n", prog);
                fprintf(stderr,"where file1 and file2 are two dna or two protein sequences.\n");
                fprintf(stderr,"The sequences can be in upper- or lower-case\n");
                fprintf(stderr,"Any lines beginning with ';' or '<' are ignored\n");
                fprintf(stderr,"Output is in the file \"align.out\"\n");
                exit(1);
        }
        namex[0] = av[1];
        namex[1] = av[2];
        seqx[0] = getseq(namex[0], &len0);
        seqx[1] = getseq(namex[1], &len1);
        xbm = (dna)? _dbval : _pbval;

        endgaps = 0;                    /* 1 to penalize endgaps */
        ofile = "align.out";            /* output file */

        nw();           /* fill in the matrix, get the possible jmps */
        readjmps();     /* get the actual jmps */
        print();        /* print stats, alignment */

        cleanup(0);     /* unlink any tmp files */
}

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()                                                            nw
{
        char            *px, *py;               /* seqs and ptrs */
        int             *ndely, *dely;          /* keep track of dely */
        int             ndelx, delx;            /* keep track of delx */
        int             *tmp;                   /* for swapping row0, row1 */
        int             mis;                    /* score for each type */
        int             ins0, ins1;             /* insertion penalties */
        register        id;                     /* diagonal index */
        register        ij;                     /* jmp index */
        register        *col0, *col1;           /* score for curr, last row */
        register        xx, yy;                 /* index into seqs */

        dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));
```

Table 1 (cont')

```
ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
ins0 = (dna)? DINS0 : PINS0;
ins1 = (dna)? DINS1 : PINS1;

smax = -10000;
if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
                col0[yy] = dely[yy] = col0[yy-1] - ins1;
                ndely[yy] = yy;
        }
        col0[0] = 0;        /* Waterman Bull Math Biol 84 */
}
else
        for (yy = 1; yy <= len1; yy++)
                dely[yy] = -ins0;

/* fill in match matrix
 */
for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
                if (xx == 1)
                        col1[0] = delx = -(ins0+ins1);
                else
                        col1[0] = delx = col0[0] - ins1;
                ndelx = xx;
        }
        else {
                col1[0] = 0;
                delx = -ins0;
                ndelx = 0;
        }

                                                                        ...nw
        for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
                mis = col0[yy-1];
                if (dna)
                        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
                else
                        mis += _day[*px-'A'][*py-'A'];

                /* update penalty for del in x seq;
                 * favor new del over ongong del
                 * ignore MAXGAP if weighting endgaps
                 */
                if (endgaps || ndely[yy] < MAXGAP) {
                        if (col0[yy] - ins0 >= dely[yy]) {
                                dely[yy] = col0[yy] - (ins0+ins1);
                                ndely[yy] = 1;
                        } else {
                                dely[yy] -= ins1;
                                ndely[yy]++;
                        }
                } else {
```

Table 1 (cont')

```
                if (col0[yy] - (ins0+ins1) >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else
                        ndely[yy]++;
        }

        /* update penalty for del in y seq;
         * favor new del over ongong del
         */
        if (endgaps || ndelx < MAXGAP) {
                if (col1[yy-1] - ins0 >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else {
                        delx -= ins1;
                        ndelx++;
                }
        } else {
                if (col1[yy-1] - (ins0+ins1) >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else
                        ndelx++;
        }

        /* pick the maximum score; we're favoring
         * mis over any del and delx over dely
         */
                                                                        ...nw
        id = xx - yy + len1 - 1;
        if (mis >= delx && mis >= dely[yy])
                col1[yy] = mis;
        else if (delx >= dely[yy]) {
                col1[yy] = delx;
                ij = dx[id].ijmp;
                if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
                && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                        dx[id].ijmp++;
                        if (++ij >= MAXJMP) {
                                writejmps(id);
                                ij = dx[id].ijmp = 0;
                                dx[id].offset = offset;
                                offset += sizeof(struct jmp) + sizeof(offset);
                        }
                }
                dx[id].jp.n[ij] = ndelx;
                dx[id].jp.x[ij] = xx;
                dx[id].score = delx;
        }
        else {
                col1[yy] = dely[yy];
                ij = dx[id].ijmp;

if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
                && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
```

Table 1 (cont')

```
                                        dx[id].ijmp++;

                                        if (++ij >= MAXJMP) {
                                                writejmps(id);
                                                ij = dx[id].ijmp = 0;
                                                dx[id].offset = offset;
                                                offset += sizeof(struct jmp) + sizeof(offset);
                                        }
                                }
                                dx[id].jp.n[ij] = -ndely[yy];
                                dx[id].jp.x[ij] = xx;
                                dx[id].score = dely[yy];
                        }
                        if (xx == len0 && yy < len1) {
                                /* last col
                                 */
                                if (endgaps)
                                        col1[yy] -= ins0+ins1*(len1-yy);
                                if (col1[yy] > smax) {
                                        smax = col1[yy];
                                        dmax = id;
                                }
                        }
                }
                if (endgaps && xx < len0)
                        col1[yy-1] -= ins0+ins1*(len0-xx);
                if (col1[yy-1] > smax) {
                        smax = col1[yy-1];
                        dmax = id;
                }
                tmp = col0; col0 = col1; col1 = tmp;
        }
        (void) free((char *)ndely);
        (void) free((char *)dely);
        (void) free((char *)col0);
        (void) free((char *)col1);
}
/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[ ]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC     3
#define P_LINE  256     /* maximum output line */
#define P_SPC   3       /* space between name or num and seq */
```

Table 1 (cont')

```
extern  _day[26][26];

int     olen;           /* set output line length */

FILE    *fx;            /* output file */

print()                                                         print
{
        int     lx, ly, firstgap, lastgap;      /* overlap */

        if ((fx = fopen(ofile, "w")) == 0) {
                fprintf(stderr,"%s: can't write %s\n", prog, ofile);
                cleanup(1);
        }
        fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
        fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
        olen = 60;
        lx = len0;
        ly = len1;
        firstgap = lastgap = 0;
        if (dmax < len1 - 1) {      /* leading gap in x */
                pp[0].spc = firstgap = len1 - dmax - 1;
                ly -= pp[0].spc;
        }
        else if (dmax > len1 - 1) { /* leading gap in y */
                pp[1].spc = firstgap = dmax - (len1 - 1);
                lx -= pp[1].spc;
        }
        if (dmax0 < len0 - 1) {     /* trailing gap in x */
                lastgap = len0 - dmax0 -1;
                lx -= lastgap;
        }
        else if (dmax0 > len0 - 1) {        /* trailing gap in y */
                lastgap = dmax0 - (len0 - 1);
                ly -= lastgap;
        }
        getmat(lx, ly, firstgap, lastgap);
        pr_align();
}

/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                               getmat
        int     lx, ly;                 /* "core" (minus endgaps) */
        int     firstgap, lastgap;      /* leading trailing overlap */
{
        int             nm, i0, i1, siz0, siz1;
        char            outx[32];
        double          pct;
        register        n0, n1;
        register char   *p0, *p1;

        /* get total matches, score
         */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
```

```
p1 = seqx[1] + pp[0].spc;
```

Table 1 (cont')

```
n0 = pp[1].spc + 1;
n1 = pp[0].spc + 1;

nm = 0;
while ( *p0 && *p1 ) {
        if (siz0) {
                p1++;
                n1++;
                siz0--;
        }
        else if (siz1) {
                p0++;
                n0++;
                siz1--;
        }
        else {
                if (xbm[*p0-'A']&xbm[*p1-'A'])
                        nm++;
                if (n0++ == pp[0].x[i0])
                        siz0 = pp[0].n[i0++];
                if (n1++ == pp[1].x[i1])
                        siz1 = pp[1].n[i1++];
                p0++;
                p1++;
        }
}

/* pct homology:
 * if penalizing endgaps, base is the shorter seq
 * else, knock off overhangs and take shorter core
 */
if (endgaps)
        lx = (len0 < len1)? len0 : len1;
else
        lx = (lx < ly)? lx : ly;
pct = 100.*(double)nm/(double)lx;
fprintf(fx, "\n");
fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity\n",
                        nm, (nm == 1)? "" : "es", lx, pct);

fprintf(fx, "<gaps in first sequence: %d", gapx);          ...getmat
if (gapx) {
        (void) sprintf(outx, " (%d %s%s)",
                ngapx, (dna)? "base":"residue", (ngapx == 1)? "":"s");
        fprintf(fx,"%s", outx);

fprintf(fx, ", gaps in second sequence: %d", gapy);
if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
                ngapy, (dna)? "base":"residue", (ngapy == 1)? "":"s");
        fprintf(fx,"%s", outx);
}
if (dna)
        fprintf(fx,
        "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
        smax, DMAT, DMIS, DINS0, DINS1);
```

Table 1 (cont')

```
        else
                fprintf(fx,
                "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
                smax, PINS0, PINS1);

        if (endgaps)
                fprintf(fx,
                "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
                firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
                lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
        else
                fprintf(fx, "<endgaps not penalized\n");
}
static          nm;             /* matches in core -- for checking */
static          lmax;           /* lengths of stripped file names */
static          ij[2];          /* jmp index for a path */
static          nc[2];          /* number at start of current line */
static          ni[2];          /* current elem number -- for gapping */
static          siz[2];
static char     *ps[2];         /* ptr to current element */
static char     *po[2];         /* ptr to next output char slot */
static char     out[2][P_LINE]; /* output line */
static char     star[P_LINE];   /* set by stars() */

/*
 * print alignment of described in struct path pp[ ]
 */
static
pr_align()                                                      pr_align
{
        int             nn;     /* char count */
        int             more;
        register        i;

        for (i =  0, lmax = 0; i < 2; i++) {
                nn = stripname(namex[i]);
                if (nn > lmax)
                        lmax = nn;

                nc[i] = 1;
                ni[i] = 1;
                siz[i] = ij[i] = 0;
                ps[i] = seqx[i];
                po[i] = out[i];
        }
        for (nn = nm = 0, more = 1; more; ) {                   ...pr_align
                for (i = more = 0; i < 2; i++) {
                        /*
                         * do we have more of this sequence?
                         */
                        if (!*ps[i])
                                continue;
                        more++;

                        if (pp[i].spc) {    /* leading space */
                                *po[i]++ = ' ';
                                pp[i].spc--;
```

Table 1 (cont')

```
                }
                else if (siz[i]) {    /* in a gap */
                        *po[i]++ = '-';
                        siz[i]--;
                }
                else {                /* we're putting a seq element
                                       */

                        *po[i] = *ps[i];
                        if (islower(*ps[i]))
                                *ps[i] = toupper(*ps[i]);
                        po[i]++;
                        ps[i]++;

                        /*
                         * are we at next gap for this seq?
                         */
                        if (ni[i] == pp[i].x[ij[i]]) {
                                /*
                                 * we need to merge all gaps
                                 * at this location
                                 */
                                siz[i] = pp[i].n[ij[i]++];
                                while (ni[i] == pp[i].x[ij[i]])
                                        siz[i] += pp[i].n[ij[i]++];
                        }
                        ni[i]++;
                }
        }
        if (++nn == olen || !more && nn) {
                dumpblock();
                for (i = 0; i < 2; i++)
                        po[i] = out[i];
                nn = 0;
        }
    }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()                                                    dumpblock
{
    register i;

    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';                                       ...dumpblock

    (void) putc('\n', fx);
    for (i = 0; i < 2; i++) {
        if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
            if (i == 0)
                nums(i);
            if (i == 0 && *out[1])
                stars();
            putline(i);
```

Table 1 (cont')

```
                        if (i == 0 && *out[1])
                                fprintf(fx, star);
                        if (i == 1)
                                nums(i);
                }
        }
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)
        nums
        int     ix;     /* index in out[ ] holding seq line */
{
        char            nline[P_LINE];
        register        i, j;
        register char   *pn, *px, *py;
        for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
                *pn = ' ';
        for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
                if (*py == ' ' || *py == '-')
                        *pn = ' ';
                else {
                        if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                                j = (i < 0)? -i : i;
                                for (px = pn; j; j /= 10, px--)
                                        *px = j%10 + '0';
                                if (i < 0)
                                        *px = '-';
                        }
                        else
                                *pn = ' ';
                        i++;
                }
        }
        *pn = '\0';
        nc[ix] = i;
        for (pn = nline; *pn; pn++)
                (void) putc(*pn, fx);
        (void) putc('\n', fx);
}
/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)                                                    putline
        int     ix;
{


                                                              ...putline
        int             i;
        register char   *px;

        for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
                (void) putc(*px, fx);
```

Table 1 (cont')

```
        for (; i < lmax+P_SPC; i++)
                (void) putc(' ', fx);

        /* these count from 1:
         * ni[ ] is current element (from 1)
         * nc[ ] is number at start of current line
         */
        for (px = out[ix]; *px; px++)
                (void) putc(*px&0x7F, fx);
        (void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()                                                         stars
{
        int             i;
        register char   *p0, *p1, cx, *px;

        if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
           !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
                return;
        px = star;
        for (i = lmax+P_SPC; i; i--)
                *px++ = ' ';

        for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
                if (isalpha(*p0) && isalpha(*p1)) {

                        if (xbm[*p0-'A']&xbm[*p1-'A']) {
                                cx = '*';
                                nm++;
                        }
                        else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                                cx = '.';
                        else
                                cx = ' ';
                }
                else
                        cx = ' ';
                *px++ = cx;
        }
        *px++ = '\n';
        *px = '\0';
}

/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)                                                   stripname
        char    *pn;    /* file name (may be path) */
{
        register char   *px, *py;
```

Table 1 (cont')

```
        py = 0;
        for (px = pn; *px; px++)
                if (*px == '/')
                        py = px + 1;
        if (py)
                (void) strcpy(pn, py);
        return(strlen(pn));

}
/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";          /* tmp file for jmps */
FILE    *fj;

int     cleanup();                           /* cleanup tmp file */
long    lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                                          cleanup
        int     i;
{
        if (fj)
                (void) unlink(jname);
        exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                                                   getseq
        char    *file;  /* file name */
        int     *len;   /* seq len */
{
        char            line[1024], *pseq;
        register char   *px, *py;
        int             natgc, tlen;
        FILE            *fp;

        if ((fp = fopen(file,"r")) == 0) {
                fprintf(stderr,"%s: can't read %s\n", prog, file);
                exit(1);
        }
        tlen = natgc = 0;
```

Table 1 (cont')

```
        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++)
                        if (isupper(*px) || islower(*px))
                                tlen++;
        }
        if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
                fprintf(stderr,"%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
                exit(1);
        }
                                pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
                                                                                ...getseq
        py = pseq + 4;
        *len = tlen;
        rewind(fp);

        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++) {
                        if (isupper(*px))
                                *py++ = *px;
                        else if (islower(*px))
                                *py++ = toupper(*px);
                        if (index("ATGCU",*(py-1)))
                                natgc++;
                }
        }
        *py++ = '\0';
        *py = '\0';
        (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
}

char    *
g_calloc(msg, nx, sz)                                                   g_calloc
        char    *msg;           /* program, calling routine */
        int     nx, sz;         /* number and size of elements */
{
        char            *px, *calloc();
        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
                if (*msg) {
                        fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                        exit(1);
                }
        }
        return(px);
}

/*
 * get final jmps from dx[ ] or tmp file, set pp[ ], reset dmax: main()
 */
readjmps()                                                              readjmps
{
        int             fd = -1;
```

Table 1 (cont')

```
        int             siz, i0, i1;
        register i, j, xx;

        if (fj) {
                (void) fclose(fj);
                if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                        fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                        cleanup(1);
                }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
                while (1) {
                        for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                                ;
                                                                                ...readjmps
                        if (j < 0 && dx[dmax].offset && fj) {
                                (void) lseek(fd, dx[dmax].offset, 0);
                                (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
                                (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
                                dx[dmax].ijmp = MAXJMP-1;
                        }
                        else
                                break;
                }
                if (i >= JMPS) {
                        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
                        cleanup(1);
                }
                if (j >= 0) {
                        siz = dx[dmax].jp.n[j];
                        xx = dx[dmax].jp.x[j];
                        dmax += siz;
                        if (siz < 0) {          /* gap in second seq */
                                pp[1].n[i1] = -siz;
                                xx += siz;

                                /* id = xx - yy + len1 - 1
                                 */
                                pp[1].x[i1] = xx - dmax + len1 - 1;
                                gapy++;
                                ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
                                siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
                                i1++;
                        }
                        else if (siz > 0) {  /* gap in first seq */
                                pp[0].n[i0] = siz;
                                pp[0].x[i0] = xx;
                                gapx++;
                                ngapx += siz;
/* ignore MAXGAP when doing endgaps */
                                siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
                                i0++;
                        }
                }
                else
                        break;
```

Table 1 (cont')

```
        }
        /* reverse the order of jmps
         */
        for (j = 0, i0--; j < i0; j++, i0--) {
                i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
                i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
        }
        for (j = 0, i1--; j < i1; j++, i1--) {
                i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
                i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
        }
        if (fd >= 0)
                (void) close(fd);
        if (fj) {
                (void) unlink(jname);
                fj = 0;
                offset = 0;
        }
}

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(ix)                                                    writejmps
        int     ix;
{
        char    *mktemp();

        if (!fj) {
                if (mktemp(jname) < 0) {
                        fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                        cleanup(1);
                }
                if ((fj = fopen(jname, "w")) == 0) {
                        fprintf(stderr, "%s: can't write %s\n", prog, jname);
                        exit(1);
                }
        }
        (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
        }
```

For purposes herein, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations, Tables 2A-B demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO".

Tables 2A-D show hypothetical exemplifications for using the below described method to determine % amino acid sequence identity (Tables 2A-B) and % nucleic acid sequence identity (Tables 2C-D) using the ALIGN-2 sequence comparison computer program, wherein "PRO" represents the amino acid sequence of a hypothetical PRO23203 polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, "PRO-DNA" represents a hypothetical PRO23203-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, "X, "Y" and "Z" each represent different hypothetical amino acid residues and "N", "L" and "V" each represent different hypothetical nucleotides.

Table 2A

PRO	XXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXYYYYYYY	(Length = 12 amino acids)

% amino acid sequence identity = (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) = 5 divided by 15 = 33.3%

Table 2B

PRO	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity = (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the

PRO polypeptide) = 5 divided by 10 = 50%

Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described above using the ALIGN-2 sequence comparison computer program. However, % amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from http://www.ncbi.nlm.nih.gov or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

"PRO23203 variant polynucleotide" or "PRO23203 variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO23203 polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with either (a) a nucleic acid sequence which encodes residues 1 to 454 of the PRO23203 polypeptide shown in Figure 2 (SEQ ID NO:2), (b) a nucleic acid sequence which encodes amino acids 1 to X of Figure 2 (SEQ ID NO:2), wherein X is any amino acid from amino acid 205 to 214 of Figure 2 (SEQ ID NO:2) or (c) a nucleic acid sequence which encodes another specifically derived fragment of the amino acid sequence shown in Figure 2 (SEQ ID NO:2). Ordinarily, a PRO23203 variant

polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with either (a) a nucleic acid sequence which encodes residues 1 to 454 of the PRO23203 polypeptide shown in Figure 2 (SEQ ID NO:2), (b) a nucleic acid sequence which encodes amino acids 1 to X of Figure 2 (SEQ ID NO:2), wherein X is any amino acid from amino acid 205 to 214 of Figure 2 (SEQ ID NO:2) or (c) a nucleic acid sequence which encodes another specifically derived fragment of the amino acid sequence shown in Figure 2 (SEQ ID NO:2). PRO23203 polynucleotide variants do not encompass the native PRO23203 nucleotide sequence.

Ordinarily, PRO23203 variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

"Percent (%) nucleic acid sequence identity" with respect to the PRO23203 polypeptide-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in a PRO23203 polypeptide-encoding nucleic acid sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of

determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared. For purposes herein, however, % nucleic acid sequence identity values are obtained as described below by using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

For purposes herein, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 2C-D demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA".

Table 2C

PRO-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity = (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) = 6 divided by 14 = 42.9%

Table 2D

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

% nucleic acid sequence identity = (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) = 4 divided by 12 = 33.3%

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described above using the ALIGN-2 sequence comparison computer program. However, % nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from http://www.ncbi.nlm.nih.gov or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

In other embodiments, PRO23203 variant polynucleotides are nucleic acid molecules that encode an active PRO23203 polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding the full-length PRO23203 polypeptide shown in Figure 2 (SEQ ID NO:2). PRO23203 variant polypeptides may be those that are encoded by a PRO23203 variant polynucleotide.

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Preferably, the isolated polypeptide is free of association with all components with which it is naturally associated. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO23203 natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" nucleic acid molecule encoding a PRO23203 polypeptide is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO23203-encoding nucleic acid. Preferably, the isolated nucleic is free of association with all components with which it is naturally associated. An isolated PRO23203-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated nucleic acid molecules therefore are distinguished from the PRO23203-encoding nucleic acid molecule as it exists in natural cells. However, an isolated nucleic acid molecule encoding a PRO23203 polypeptide includes PRO23203-encoding nucleic acid molecules contained in cells that ordinarily express PRO23203 where, for example, the nucleic acid molecule is in a

chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO23203 monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO23203 antibody compositions with polyepitopic specificity, single chain anti-PRO23203 antibodies, and fragments of anti-PRO23203 antibodies (see below). The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

The monoclonal antibodies herein include "chimeric" antibodies in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity (see U.S. Patent No. 4,816,567; and Morrison *et al., Proc. Natl. Acad. Sci. USA*, 81:6851-6855 (1984)). Chimeric antibodies of interest herein include "primatized" antibodies comprising variable domain antigen-binding sequences derived from

a non-human primate (*e.g.* Old World Monkey, Ape etc), and human constant region sequences.

An "intact" antibody is one which comprises an antigen-binding site as well as a C_L and at least heavy chain constant domains, C_H1, C_H2 and C_H3. The constant domains may be native sequence constant domains (*e.g.* human native sequence constant domains) or amino acid sequence variant thereof. Preferably, the intact antibody has one or more effector functions.

As used herein, an anti-PUMPCn antibody that "internalizes" is one that is taken up by (i.e., enters) the cell upon binding to PUMPCn on a mammalian cell (i.e. cell surface PUMPCn). The internalizing antibody will of course include antibody fragments, human or humanized antibody and antibody conjugate. For therapeutic applications, internalization in vivo is contemplated. The number of antibody molecules internalized will be sufficient or adequate to kill a PUMPCn-expressing cell, especially a PUMPCn-expressing cancer cell. Depending on the potency of the antibody or antibody conjugate, in some instances, the uptake of a single antibody molecule into the cell is sufficient to kill the target cell to which the antibody binds. For example, certain toxins are highly potent in killing such that internalization of one molecule of the toxin conjugate to the antibody is sufficient to kill the tumor cell.

Whether an anti-PUMPCn antibody internalizes upon binding PUMPCn on a mammalian cell can be determined by various assays. For example, to test internalization in vivo, the test antibody is labeled and introduced into an animal known to have PUMPCn expressed on the surface of certain cells. The antibody can be radiolabeled or labeled with fluorescent or gold particles, for instance. Animals suitable for this assay include a mammal such as a NCR nude mouse that contains a human PUMPCn-expressing tumor transplant or xenograft, or a mouse into which cells transfected with human PUMPCn have been introduced, or a transgenic mouse expressing the human PUMPCn transgene. Appropriate controls include animals that did not receive the test antibody or that received an unrelated antibody, and animals that received an antibody to another antigen on the cells of interest, which antibody is known to be internalized upon binding to the antigen (e.g., HERCEPTIN which binds to Her2 expressed on the human breast tumor cell line, MCF-7). The antibody can be administered to the animal, e.g., by intravenous injection. At suitable time intervals, tissue sections of the animal can be prepared using known methods or as described in the experimental examples below, and analyzed by light microscopy or electron microscopy, for internalization as well as the location of the internalized antibody in the cell. For

internalization *in vitro*, the cells can be incubated in tissue culture dishes in the presence or absence of the relevant antibodies added to the culture media and processed for microscopic analysis at desired time points. The presence of an internalized, labeled antibody in the cells can be directly visualized by microscopy or by autoradiography if radiolabeled antibody is used. Alternatively, in a quantitative biochemical assay, a population of cells comprising PUMPCn-expressing cells are contacted *in vitro* or *in vivo* with a radiolabeled test antibody and the cells (if contacted *in vivo*, cells are then isolated after a suitable amount of time) are treated with a protease or subjected to an acid wash to remove uninternalized antibody on the cell surface. The cells are ground up and the amount of protease resistant, radioactive counts per minute (cpm) associated with each batch of cells is measured by passing the homogenate through a scintillation counter. Based on the known specific activity of the radiolabeled antibody, the number of antibody molecules internalized per cell can be deduced from the scintillation counts of the ground-up cells. Cells are "contacted" with antibody *in vitro* preferably in solution form such as by adding the cells to the cell culture media in the culture dish or flask and mixing the antibody well with the media to ensure uniform exposure of the cells to the antibody. Instead of adding to the culture media, the cells can be contacted with the test antibody in an isotonic solution such as PBS in a test tube for the desired time period. *In vivo*, the cells are contacted with antibody by any suitable method of administering the test antibody such as the methods of administration described below when administered to a patient.

The faster the rate of internalization of the antibody upon binding to the PUMPCn expressing cell *in vivo*, the faster the desired killing or growth inhibitory effect on the target PUMPCn-expressing cell can be achieved, e.g., by a cytotoxic immunoconjugate. Preferably, the kinetics of internalization of the anti-PUMPCn antibodies are such that they favor rapid killing of the PUMPCn-expressing target cell. Therefore, it is desirable that the anti-PUMPCn antibody exhibit a rapid rate of internalization preferably, within 24 hours from administration of the antibody *in vivo*, more preferably within about 12 hours.

To determine if a test antibody can compete for binding to the same epitope as the epitope bound by the anti-PUMPCn antibodies of the present invention, a cross-blocking assay e.g., a competitive ELISA assay can be performed. In an exemplary competitive ELISA assay, PUMPCn or the relevant fragment thereof coated on the wells of a microtiter plate is pre-incubated with or without candidate competing antibody and then the biotin-labeled anti-PUMPCn antibody of the invention is added. The amount of labeled anti-PUMPCn antibody bound to the PUMPCn antigen in the wells is measured using avidin-peroxidase conjugate and

appropriate substrate. The antibody can be labeled with a radioactive or fluorescent label or some other detectable and measurable label. The amount of labeled anti-PUMPCn antibody that bound to the antigen will have an indirect correlation to the ability of the candidate competing antibody (test antibody) to compete for binding to the same epitope, i.e., the greater the affinity of the test antibody for the same epitope, the less labeled antibody will be bound to the antigen-coated wells. A candidate competing antibody is considered an antibody that binds substantially to the same epitope or that competes for binding to the same epitope as an anti-PUMPCn antibody of the invention if the candidate antibody can block binding of the PUMPCn antibody by at least 20%, preferably by at least 20-50%, even more preferably, by at least 50% as compared to the control performed in parallel in the absence of the candidate competing antibody (but may be in the presence of a known non-competing antibody). It will be understood that variations of this assay can be performed to arrive at the same quantitative value.

An "antibody that inhibits the growth of tumor cells expressing PUMPCn" or a "growth inhibitory" antibody is one which binds to and results in measurable growth inhibition of cancer cells expressing or overexpressing PUMPCn. Preferred growth inhibitory anti-PUMPCn antibodies inhibit growth of PUMPCn-expressing tumor cells (e.g., prostate cancer cells) by greater than 20%, preferably from about 20% to about 50%, and even more preferably, by greater than 50% (*e.g.* from about 50% to about 100%) as compared to the appropriate control, the control typically being tumor cells not treated with the antibody being tested. Growth inhibition can be measured at an antibody concentration of about 0.1 to 30 μg/ml or about 0.5 nM to 200 nM in cell culture, where the growth inhibition is determined 1-10 days after exposure of the tumor cells to the antibody. The antibody is growth inhibitory in vivo if administration of the anti-PUMPCn antibody at about 1 μg/kg to about 100 mg/kg body weight results in reduction in tumor size or tumor cell proliferation within about 5 days to 3 months from the first administration of the antibody, preferably within about 5 to 30 days.

The terms "cancer" and "cancerous" refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Examples of cancer include, but are not limited to, carcinoma, lymphoma, blastoma, sarcoma, and leukemia or lymphoid malignancies. More particular examples of such cancers include squamous cell cancer (*e.g.* epithelial squamous cell cancer), lung cancer including small-cell lung cancer, non-small cell lung cancer, adenocarcinoma of the lung and squamous carcinoma of the lung, cancer of the peritoneum, hepatocellular cancer, gastric or stomach cancer including

gastrointestinal cancer, pancreatic cancer, glioblastoma, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, hepatoma, breast cancer, colon cancer, rectal cancer, colorectal cancer, endometrial or uterine carcinoma, salivary gland carcinoma, kidney or renal cancer, prostate cancer, vulval cancer, thyroid cancer, hepatic carcinoma, anal carcinoma, penile carcinoma, melanoma, multiple myeloma and B-cell lymphoma, brain , as well as head and neck cancer, and associated metastases.

A "PUMPCn-expressing cell" is a cell which expresses endogenous or transfected PUMPCn on the cell surface. A "PUMPCn-expressing cancer" is a cancer comprising cells that have PUMPCn protein present on the cell surface. A "PUMPCn-expressing cancer" produces sufficient levels of PUMPCn on the surface of cells thereof, such that an anti-PUMPCn antibody can bind thereto and have a therapeutic effect with respect to the cancer. A cancer which "overexpresses" PUMPCn is one which has significantly higher levels of PUMPCn at the cell surface thereof, compared to a noncancerous cell of the same tissue type. Such overexpression may be caused by gene amplification or by increased transcription or translation. PUMPCn overexpression may be determined in a diagnostic or prognostic assay by evaluating increased levels of the PUMPCn protein present on the surface of a cell (*e.g.* via an immunohistochemistry assay; FACS analysis). Alternatively, or additionally, one may measure levels of PUMPCn-encoding nucleic acid or mRNA in the cell, *e.g.* via fluorescent *in situ* hybridization; (FISH; see WO98/45479 published October, 1998), Southern blotting, Northern blotting, or polymerase chain reaction (PCR) techniques, such as real time quantitative PCR (RT-PCR). One may also study PUMPCn overexpression by measuring shed antigen in a biological fluid such as serum, e.g., using antibody-based assays (see also, *e.g.*, U.S. Patent No. 4,933,294 issued June 12, 1990; WO91/05264 published April 18, 1991; U.S. Patent 5,401,638 issued March 28, 1995; and Sias *et al. J. Immunol. Methods* 132: 73-80 (1990)). Aside from the above assays, various *in vivo* assays are available to the skilled practitioner. For example, one may expose cells within the body of the patient to an antibody which is optionally labeled with a detectable label, *e.g.* a radioactive isotope, and binding of the antibody to cells in the patient can be evaluated, *e.g.* by external scanning for radioactivity or by analyzing a biopsy taken from a patient previously exposed to the antibody. A PUMPCn-expressing cancer includes prostate and lung cancer.

"Treating" or "treatment" or "alleviation" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include

those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented. A subject or mammal is successfully "treated" for a PUMPCn-expressing cancer if, after receiving a therapeutic amount of an anti-PUMPCn antibody according to the methods of the present invention, the patient shows observable and/or measurable reduction in or absence of one or more of the following: reduction of PSA levels, reduction in the number of cancer cells or absence of the cancer cells; reduction in the tumor size; inhibition (*i.e.*, slow to some extent and preferably stop) of cancer cell infiltration into peripheral organs including the spread of cancer into soft tissue and bone; inhibition (*i.e.*, slow to some extent and preferably stop) of tumor metastasis; inhibition, to some extent, of tumor growth; and/or relief to some extent, one or more of the symptoms associated with the specific cancer; reduced morbidity and mortality, and improvement in quality of life issues. To the extent the anti-PUMPCn antibody may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic. Reduction of these signs or symptoms may also be felt by the patient.

The above parameters for assessing successful treatment and improvement in the disease are readily measurable by routine procedures familiar to a physician. For cancer therapy, efficacy can be measured, for example, by assessing the time to disease progression (TTP) and/or determining the response rate (RR). For prostate cancer, the progress of therapy can be assessed by routine methods, usually by measuring serum PSA (prostate specific antigen) levels; the higher the level of PSA in the blood, the more extensive the cancer. Commercial assays for detecting PSA are available, e.g, Hybitech Tandem-E and Tandem-R PSA assay kits, the Yang ProsCheck polyclonal assay (Yang Labs, Bellevue, WA), Abbott Imx (Abbott Labs, Abbott Park, IL), etc. Metastasis can be determined by staging tests and by bone scan and tests for calcium level and other enzymes to determine spread to the bone. CT scans can also be done to look for spread to the pelvis and lymph nodes in the area. Chest X-rays and measurement of liver enzyme levels by known methods are used to look for metastasis to the lungs and liver, respectively. Other routine methods for monitoring the disease include transrectal ultrasonography (TRUS) and transrectal needle biopsy (TRNB).

The term "therapeutically effective amount" refers to an amount of an antibody or a drug effective to "treat" a disease or disorder in a subject or mammal. In the case of cancer, the therapeutically effective amount of the drug may reduce the number of cancer cells; reduce the tumor size; inhibit (*i.e.*, slow to some extent and preferably stop) cancer cell infiltration into peripheral organs; inhibit (*i.e.*, slow to some extent and preferably stop)

tumor metastasis; inhibit, to some extent, tumor growth; and/or relieve to some extent one or more of the symptoms associated with the cancer. See preceding definition of "treating". To the extent the drug may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic.

A "growth inhibitory agent" when used herein refers to a compound or composition which inhibits growth of a cell, especially a PUMPCn expressing cancer cell, either *in vitro* or *in vivo*. Thus, the growth inhibitory agent may be one which significantly reduces the percentage of PUMPCn expressing cells in S phase. Examples of growth inhibitory agents include agents that block cell cycle progression (at a place other than S phase), such as agents that induce G1 arrest and M-phase arrest. Classical M-phase blockers include the vincas (vincristine and vinblastine), taxanes, and topoisomerase II inhibitors such as doxorubicin, epirubicin, daunorubicin, etoposide, and bleomycin. Those agents that arrest G1 also spill over into S-phase arrest, for example, DNA alkylating agents such as tamoxifen, prednisone, dacarbazine, mechlorethamine, cisplatin, methotrexate, 5-fluorouracil, and ara-C. Further information can be found in *The Molecular Basis of Cancer*, Mendelsohn and Israel, eds., Chapter 1, entitled "Cell cycle regulation, oncogenes, and antineoplastic drugs" by Murakami *et al.* (WB Saunders: Philadelphia, 1995), especially p. 13. The taxanes (paclitaxel and docetaxel) are anticancer drugs both derived from the yew tree. Docetaxel (TAXOTERE®, Rhone-Poulenc Rorer), derived from the European yew, is a semisynthetic analogue of paclitaxel (TAXOL®, Bristol-Myers Squibb). Paclitaxel and docetaxel promote the assembly of microtubules from tubulin dimers and stabilize microtubules by preventing depolymerization, which results in the inhibition of mitosis in cells.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and % SDS) less stringent that those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "overexpression" as used herein refers to overexpression of a gene and/or its encoded protein in a cell, such as a cancer cell. A cancer cell that "overexpresses" a protein is one that has significantly higher levels of that protein compared to a noncancerous cell of the same tissue type.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO23203 polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

"Active" or "activity" for the purposes herein refers to form(s) of PRO23203 which retain a biological and/or an immunological activity of native or naturally-occurring PRO23203, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO23203 other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO23203 and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO23203. A preferred biological activity includes, for example, the growth and differentiation of cells from prostate tissue.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO23203 polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO23203 polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO23203 polypeptides, peptides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO23203 polypeptide may comprise contacting a PRO23203 polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO23203 polypeptide.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

The term "effective amount" is a concentration or amount of a PRO polypeptide and/or agonist/antagonist which results in achieving a particular stated purpose. An "effective amount" of a PRO polypeptide or agonist or antagonist thereof may be determined empirically. Furthermore, a "therapeutically effective amount" is a concentration or amount of a PRO polypeptide and/or agonist/antagonist which is effective for achieving a stated therapeutic effect. This amount may also be determined empirically.

A "chemotherapeutic agent" is a chemical compound useful in the treatment of cancer. Examples of chemotherapeutic agents include alkylating agents such as thiotepa and cyclosphosphamide (CYTOXAN™); alkyl sulfonates such as busulfan, improsulfan and piposulfan; aziridines such as benzodopa, carboquone, meturedopa, and uredopa; ethylenimines and methylamelamines including altretamine, triethylenemelamine, trietylenephosphoramide, triethylenethiophosphaoramide and trimethylolomelamine; nitrogen mustards such as chlorambucil, chlornaphazine, cholophosphamide, estramustine, ifosfamide, mechlorethamine, mechlorethamine oxide hydrochloride, melphalan, novembichin, phenesterine, prednimustine, trofosfamide, uracil mustard; nitrosureas such as carmustine, chlorozotocin, fotemustine, lomustine, nimustine, ranimustine; antibiotics such as aclacinomysins, actinomycin, authramycin, azaserine, bleomycins, cactinomycin, calicheamicin, carabicin, carminomycin, carzinophilin, chromomycins, dactinomycin, daunorubicin, detorubicin, 6-diazo-5-oxo-L-norleucine, doxorubicin, epirubicin, esorubicin, idarubicin, marcellomycin, mitomycins, mycophenolic acid, nogalamycin, olivomycins, peplomycin, potfiromycin, puromycin, quelamycin, rodorubicin, streptonigrin, streptozocin, tubercidin, ubenimex, zinostatin, zorubicin; anti-metabolites such as methotrexate and 5-fluorouracil (5-FU); folic acid analogues such as denopterin, methotrexate, pteropterin, trimetrexate; purine analogs such as fludarabine, 6-mercaptopurine, thiamiprine, thioguanine; pyrimidine analogs such as ancitabine, azacitidine, 6-azauridine, carmofur, cytarabine, dideoxyuridine, doxifluridine, enocitabine, floxuridine, 5-FU; androgens such as calusterone, dromostanolone propionate, epitiostanol, mepitiostane, testolactone; anti-adrenals such as aminoglutethimide, mitotane, trilostane; folic acid replenisher such as frolinic acid; aceglatone; aldophosphamide glycoside; aminolevulinic acid; amsacrine; bestrabucil; bisantrene;

edatraxate; defofamine; demecolcine; diaziquone; elfornithine; elliptinium acetate; etoglucid; gallium nitrate; hydroxyurea; lentinan; lonidamine; mitoguazone; mitoxantrone; mopidamol; nitracrine; pentostatin; phenamet; pirarubicin; podophyllinic acid; 2-ethylhydrazide; procarbazine; PSK®; razoxane; sizofiran; spirogermanium; tenuazonic acid; triaziquone; 2, 2',2"-trichlorotriethylamine; urethan; vindesine; dacarbazine; mannomustine; mitobronitol; mitolactol; pipobroman; gacytosine; arabinoside ("Ara-C"); cyclophosphamide; thiotepa; taxanes, *e.g.* paclitaxel (TAXOL®, Bristol-Myers Squibb Oncology, Princeton, NJ) and doxetaxel (TAXOTERE®, Rhône-Poulenc Rorer, Antony, France); chlorambucil; gemcitabine; 6-thioguanine; mercaptopurine; methotrexate; platinum analogs such as cisplatin and carboplatin; vinblastine; platinum; etoposide (VP-16); ifosfamide; mitomycin C; mitoxantrone; vincristine; vinorelbine; navelbine; novantrone; teniposide; daunomycin; aminopterin; xeloda; ibandronate; CPT-11; topoisomerase inhibitor RFS 2000; difluoromethylornithine (DMFO); retinoic acid; esperamicins; capecitabine; and pharmaceutically acceptable salts, acids or derivatives of any of the above. Also included in this definition are anti-hormonal agents that act to regulate or inhibit hormone action on tumors such as anti-estrogens including for example tamoxifen, raloxifene, aromatase inhibiting 4(5)-imidazoles, 4-hydroxytamoxifen, trioxifene, keoxifene, LY117018, onapristone, and toremifene (Fareston); and anti-androgens such as flutamide, nilutamide, bicalutamide, leuprolide, and goserelin; and pharmaceutically acceptable salts, acids or derivatives of any of the above.

The term "cytotoxic agent" as used herein refers to a substance that inhibits or prevents the function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (*e.g.* At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32} and radioactive isotopes of Lu), chemotherapeutic agents, and toxins such as small molecule toxins or enzymatically active toxins of bacterial, fungal, plant or animal origin, including fragments and/or variants thereof.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other

carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', $F(ab')_2$, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an $F(ab')_2$ fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the VH-VL dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. $F(ab')_2$ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

"Single-chain Fv" or "sFv" antibody fragments comprise the VH and VL domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the VH and VL domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenburg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (VH) connected to a light-chain variable domain (VL) in the same polypeptide chain (VH - VL). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

An antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO23203 polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "microarray" is defined either as a substrate with individual DNA probes in a specifically arrayed sequence and is hybridized with labeled nucleic acids that originated in a tissue sample or a substrate with individual tissues in a specifically arrayed sequence and is hybridized with a labeled nucleic acid probe.

A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

In some patients prostate cancer is treated by androgen deprivation (e.g., hormonal therapy). If the patient's prostate cancer becomes resistant to the hormonal therapy the cancer is defined as being "Androgen independant prostate cancer".

II. Compositions and Methods of the Invention

A. Full-length PRO23203 Polypeptide

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO23203 (or also UNQ6507). In particular, cDNA encoding a PRO23203 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by DNA185171-2994 as well as all further native homologues and variants included in the foregoing definition of PRO23203, will be referred to as "PRO23203", regardless of their origin or mode of preparation.

As disclosed in the Examples below, a cDNA clone designated herein as DNA185171-

2994 has been deposited with the ATCC. The actual nucleotide sequence of the clone can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO23203 polypeptide and encoding nucleic acid described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

B. PRO23203 Variants

In addition to the full-length native sequence PRO23203 polypeptides described herein, it is contemplated that PRO23203 variants can be prepared. PRO23203 variants can be prepared by introducing appropriate nucleotide changes into the PRO23203 DNA, and/or by synthesis of the desired PRO23203 polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO23203, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO23203 or in various domains of the PRO23203 described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO23203 that results in a change in the amino acid sequence of the PRO23203 as compared with the native sequence PRO23203. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO23203. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO23203 with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

PRO23203 polypeptide fragments are provided herein. Such fragments may be

truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO23203 polypeptide.

PRO23203 fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO23203 fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO23203 polypeptide fragments share at least one biological and/or immunological activity with the native PRO23203 polypeptide shown in Figure 2 (SEQ ID NO:2).

In particular embodiments, conservative substitutions of interest are shown in Table 3 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 3, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 3

Original Residue	Exemplary Substitutions	Preferred Substitutions
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

Substantial modifications in function or immunological identity of the PRO23203 polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

(1) hydrophobic: norleucine, met, ala, val, leu, ile;

(2) neutral hydrophilic: cys, ser, thr;

(3) acidic: asp, glu;

(4) basic: asn, gln, his, lys, arg;

(5) residues that influence chain orientation: gly, pro; and

(6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO23203 variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

C. Modifications of PRO23203

Covalent modifications of PRO23203 are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO23203 polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO23203. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO23203 to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO23203 antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α-amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure

and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO23203 polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO23203 (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO23203. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO23203 polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO23203 (for O-linked glycosylation sites). The PRO23203 amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO23203 polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO23203 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO23203 polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO23203 comprises linking the PRO23203 polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO23203 of the present invention may also be modified in a way to form a chimeric molecule comprising PRO23203 fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the PRO23203 with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO23203. The presence of such epitope-tagged forms of the PRO23203 can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO23203 to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α-tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO23203 with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Immunoglobulin fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO23203 polypeptide in place of at least one variable region within an Immunoglobulin molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

D. Preparation of PRO23203

The description below relates primarily to production of PRO23203 by culturing cells transformed or transfected with a vector containing PRO23203 nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to

prepare PRO23203. For instance, the PRO23203 sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO23203 may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO23203.

1. Isolation of DNA Encoding PRO23203

DNA encoding PRO23203 may be obtained from a cDNA library prepared from tissue believed to possess the PRO23203 mRNA and to express it at a detectable level. Accordingly, human PRO23203 DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO23203-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as antibodies to the PRO23203 or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO23203 is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ^{32}P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level)

within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO23203 production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, $CaCl_2$, $CaPO_4$, liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA* ; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan*^{r}; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan*^{r}; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO23203-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilarum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus; yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida; Trichoderma reesia* (EP

244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora, Penicillium, Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Balance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylotropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula, Candida, Kloeckera, Pichia, Saccharomyces, Torulopsis,* and *Rhodotorula.* A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Suitable host cells for the expression of glycosylated PRO23203 are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as Drosophila S2 and Spodoptera Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO23203 may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The PRO23203 may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO23203-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α-factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO23203-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp*1 gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp*1 gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO23203-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β-lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO23203.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO23203 transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the PRO23203 by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic

cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5′ or 3′ to the PRO23203 coding sequence, but is preferably located at a site 5′ from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5′ and, occasionally 3′, untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO23203.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO23203 in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out here the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO23203 polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO23203 DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO23203 may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO23203 can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO23203 from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO23203. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO23203 produced.

E. Uses for PRO23203

Nucleotide sequences (or their complement) encoding PRO23203 have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO23203 nucleic acid will also be useful for the preparation of PRO23203 polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO23203 gene (SEQ ID NO:1), or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO23203 cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO23203 or PRO23203 from other species) which have a desired sequence identity to the PRO23203 sequence disclosed in Figure 1 (SEQ ID NO:1). Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the nucleotide sequence of SEQ ID NO:1 wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO23203. By way of example, a screening method will comprise isolating the coding region of the PRO23203 gene using the

known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO23203 gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

Other useful fragments of the PRO23203 nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO23203 mRNA (sense) or PRO23203 DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO23203 DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO23203 proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide

sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example, $CaPO_4$-mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO23203 coding sequences.

Nucleotide sequences encoding a PRO23203 can also be used to construct hybridization probes for mapping the gene which encodes that PRO23203 and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO23203 encode a protein which binds to another protein (example, where the PRO23203 is a receptor), the PRO23203 can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in

such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO23203 can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO23203 or a receptor for PRO23203. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO23203 or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO23203 can be used to clone genomic DNA encoding PRO23203 in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO23203. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO23203 transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO23203 introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO23203. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO23203 can be used to construct a PRO23203 "knock out" animal which has a defective or altered gene encoding PRO23203 as a result of homologous recombination between the endogenous gene encoding PRO23203 and altered genomic DNA encoding PRO23203 introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO23203 can be used to clone genomic DNA

encoding PRO23203 in accordance with established techniques. A portion of the genomic DNA encoding PRO23203 can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., Cell, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO23203 polypeptide.

Nucleic acid encoding the PRO23203 polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation,

microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau et al., Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu et al., J. Biol. Chem. 262, 4429-4432 (1987); and Wagner et al., Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson et al., Science 256, 808-813 (1992).

The PRO23203 polypeptides described herein may also be employed as molecular weight markers for protein electrophoresis purposes.

The nucleic acid molecules encoding the PRO23203 polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO23203 nucleic acid molecule of the present invention can be used as a chromosome marker.

The PRO23203 polypeptides and nucleic acid molecules of the present invention may also be used for tissue typing, wherein the PRO23203 polypeptides of the present invention may be differentially expressed in one tissue as compared to another. PRO23203 nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

The PRO23203 polypeptides described herein may also be employed as therapeutic agents. The PRO23203 polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO23203 product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or

stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, PLURONICS™ or PEG.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In Toxicokinetics and New Drug Development, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

When *in vivo* administration of a PRO23203 polypeptide or agonist or antagonist thereof is employed, normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 μg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

Where sustained-release administration of a PRO23203 polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO23203 polypeptide, microencapsulation of the PRO23203 polypeptide is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN-), interleukin-2, and MN rgp120. Johnson et al., Nat. Med., 2:795-799 (1996); Yasuda, Biomed. Ther., 27:1221-1223 (1993); Hora et al., Bio/Technology, 8:755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

The sustained-release formulations of these proteins were developed using poly-lactic-coglycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

This invention encompasses methods of screening compounds to identify those that mimic the PRO23203 polypeptide (agonists) or prevent the effect of the PRO23203 polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO23203 polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

All assays for antagonists are common in that they call for contacting the drug candidate with a PRO23203 polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

In binding assays, the interaction is binding and the complex formed can be isolated or

detected in the reaction mixture. In a particular embodiment, the PRO23203 polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO23203 polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO23203 polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

If the candidate compound interacts with but does not bind to a particular PRO23203 polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β-galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein

domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO23203 polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the gene product, a PRO23203 polypeptide and its reaction partner.

To assay for antagonists, the PRO23203 polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO23203 polypeptide indicates that the compound is an antagonist to the PRO23203 polypeptide. Alternatively, antagonists may be detected by combining the PRO23203 polypeptide and a potential antagonist with membrane-bound PRO23203 polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO23203 polypeptide can be labeled, such as by radioactivity, such that the number of PRO23203 polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO23203 polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO23203 polypeptide. Transfected cells that are grown on glass slides are exposed to labeled PRO23203 polypeptide. The PRO23203 polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the

putative receptor.

As an alternative approach for receptor identification, labeled PRO23203 polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro- sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled PRO23203 polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO23203 polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO23203 polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO23203 polypeptide.

Another potential PRO23203 polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5′ coding portion of the polynucleotide sequence, which encodes the mature PRO23203 polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the PRO23203 polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO23203 polypeptide (antisense - Okano, Neurochem., 56:560 (1991);

Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO23203 polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO23203 polypeptide, thereby blocking the normal biological activity of the PRO23203 polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, Current Biology, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

Potential uses for PRO23203 polypeptides, nucleic acids encoding therefor, and anti-PRO23203 antibodies are in the detection and/or treatment of tumors of prostate tissue origin.

F. Anti-PRO23203 Antibodies

The present invention further provides anti-PRO23203 antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

The anti-PUMPCn antibodies of the invention also have various non-therapeutic applications. In view of the fact that the use of PSA as a tool for screening or diagnosing

prostate cancer is controversial, the anti-PUMPCn antibodies of the present invention can be useful for diagnosis and staging of PUMPCn-expressing cancers (e.g., in radioimaging). The antibodies are also useful for purification or immunoprecipitation of PUMPCn from cells, for detection and quantitation of PUMPCn in vitro, e.g. in an ELISA or a Western blot, to kill and eliminate PUMPCn-expressing cells from a population of mixed cells as a step in the purification of other cells.

The present anti-PUMPCn antibodies are useful for treating a PUMPCn-expressing cancer or alleviating one or more symptoms of the cancer in a mammal. The cancers encompass metastatic cancers e.g., prostate cancer metastases. The antibody is able to bind to at least a portion of the cancer cells that express PUMPCn in the mammal and preferably is one that does not induce or that minimizes HAMA response. In a preferred embodiment, the antibody is effective to destroy or kill PUMPCn-expressing tumor cells or inhibit the growth of such tumor cells, in vitro or in vivo, upon binding to PUMPCn on the cell. Such an antibody includes a naked anti-PUMPCn antibody (not conjugated to any agent). Naked antibodies that have cytotoxic or cell growth inhibition properties can be further harnessed with a cytotoxic agent to render them even more potent in tumor cell destruction. Cytotoxic properties can be conferred to an anti-PUMPCn antibody by, e.g., conjugating the antibody with a cytotoxic agent, to form an immunoconjugate as described below. The cytotoxic agent or a growth inhibitory agent is preferably a small molecule. Toxins such as calicheamicin or a maytansinoid and analogs or derivatives thereof, are preferable.

The invention provides a composition comprising an anti-PUMPCn antibody of the invention, and a carrier. For the purposes of treating cancer, compositions can be administered to the patient in need of such treatment, wherein the composition can comprise one or more anti-PUMPCn antibodies present as an immunoconjugate or as the naked antibody. In a further embodiment, the compositions can comprise these antibodies in combination with other therapeutic agents such as cytotoxic or growth inhibitory agents, including chemotherapeutic agents. The invention also provides formulations comprising an anti-PUMPCn antibody of the invention, and a carrier. In one embodiment, the formulation is a therapeutic formulation comprising a pharmaceutically acceptable carrier.

1. Polyclonal Antibodies

The anti-PRO23203 antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if

desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO23203 polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

2. Monoclonal Antibodies

The anti-PRO23203 antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO23203 polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma

lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO23203. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., supra] or by covalently joining to the immunoglobulin coding sequence all or

part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Human and Humanized Antibodies

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

The anti-PRO23203 antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(antibody')$_2$ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary

determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

The choice of human variable domains, both light and heavy, to be used in making the humanized antibodies is based in large part on the ability to reduce antigenicity and HAMA response (human anti-mouse antibody) when the antibody is intended for human therapeutic use. According to the so-called "best-fit" method, the sequence of the variable domain of a rodent antibody is screened against the entire library of known human variable domain sequences. The human V domain sequence which is closest to that of the rodent is identified and the human framework region (FR) within it accepted for the humanized antibody (Sims *et al., J. Immunol.*, 151:2296 (1993); Chothia *et al., J. Mol. Biol.*, 196:901 (1987)). Another method uses a particular framework region derived from the consensus sequence of all human antibodies of a particular subgroup of light or heavy chains. The same framework may be used for several different humanized antibodies (Carter *et al., Proc. Natl. Acad. Sci. USA*, 89:4285 (1992); Presta *et al., J. Immunol.*, 151:2623 (1993)).

Another consideration is that antibodies be humanized with retention of high binding affinity for the antigen and other favorable biological properties. To achieve this goal, according to one method, humanized antibodies are prepared by a process of analysis of the parental sequences and various conceptual humanized products using three-dimensional models of the parental and humanized sequences. Three-dimensional immunoglobulin models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three-dimensional conformational structures of selected candidate immunoglobulin sequences. Inspection of these displays permits analysis of the likely role of the residues in the functioning of the

candidate immunoglobulin sequence, *i.e.*, the analysis of residues that influence the ability of the candidate immunoglobulin to bind its antigen. In this way, FR residues can be selected and combined from the recipient and import sequences so that the desired antibody characteristic, such as increased affinity for the target antigen(s), is achieved. In general, the hypervariable region residues are directly and most substantially involved in influencing antigen binding.

Various forms of a humanized anti-PUMPCn antibody are contemplated. For example, the humanized antibody may be an antibody fragment, such as a Fab, which is optionally conjugated with one or more cytotoxic agent(s) in order to generate an immunoconjugate. Alternatively, the humanized antibody may be an intact antibody, such as an intact IgG1 antibody.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*, Bio/Technology 10, 779-783 (1992); Lonberg *et al.*, Nature 368 856-859 (1994); Morrison, Nature 368, 812-13 (1994); Fishwild *et al.*, Nature Biotechnology 14, 845-51 (1996); Neuberger, Nature Biotechnology 14, 826 (1996); Lonberg and Huszar, Intern. Rev. Immunol. 13 65-93 (1995). Human antibodies may also be generated by *in vitro* activated B cells (see U.S. Patents 5,567,610 and 5,229,275).

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO23203, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(antibody')$_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(antibody')$_2$ fragments. These fragments are reduced in the presence of the dithiol

complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby *et al.*, J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(antibody')$_2$ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny *et al.*, J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, J. Immunol. 152:5368 (1994). Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt *et al.*, J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies may bind to two different epitopes on a given PRO23203 polypeptide herein. Alternatively, an anti-PRO23203 polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as

FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO23203 polypeptide. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express a particular PRO23203 polypeptide. These antibodies possess a PRO23203-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO23203 polypeptide and further binds tissue factor (TF).

5. Multivalent Antibodies

A multivalent antibody may be internalized (and/or catabolized) faster than a bivalent antibody by a cell expressing an antigen to which the antibodies bind. The antibodies of the present invention can be multivalent antibodies (which are other than of the IgM class) with three or more antigen binding sites (e.g. tetravalent antibodies), which can be readily produced by recombinant expression of nucleic acid encoding the polypeptide chains of the antibody. The multivalent antibody can comprise a dimerization domain and three or more antigen binding sites. The preferred dimerization domain comprises (or consists of) an Fc region or a hinge region. In this scenario, the antibody will comprise an Fc region and three or more antigen binding sites amino-terminal to the Fc region. The preferred multivalent antibody herein comprises (or consists of) three to about eight, but preferably four, antigen binding sites. The multivalent antibody comprises at least one polypeptide chain (and preferably two polypeptide chains), wherein the polypeptide chain(s) comprise two or more variable domains. For instance, the polypeptide chain(s) may comprise VD1-$(X1)_n$-VD2-$(X2)_n$-Fc, wherein VD1 is a first variable domain, VD2 is a second variable domain, Fc is one polypeptide chain of an Fc region, X1 and X2 represent an amino acid or polypeptide, and n is 0 or 1. For instance, the polypeptide chain(s) may comprise: VH-CH1-flexible linker-VH-CH1-Fc region chain; or VH-CH1-VH-CH1-Fc region chain. The multivalent antibody herein preferably further comprises at least two (and preferably four) light chain variable domain polypeptides. The multivalent antibody herein may, for instance, comprise from about two to about eight light chain variable domain polypeptides. The light chain variable domain polypeptides contemplated here comprise a light chain variable domain and, optionally, further comprise a CL domain.

6. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies

have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

7. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design, 3: 219-230 (1989).

8. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include

^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is conjugated to a cytotoxic agent (*e.g.*, a radionucleotide).

Maytansine and maytansinoids

In one preferred embodiment, an anti-PUMPCn antibody (full length or fragments) of the invention is conjugated to one or more maytansinoid molecules.

Maytansinoids are mitototic inhibitors which act by inhibiting tubulin polymerization. Maytansine was first isolated from the east African shrub *Maytenus serrata* (U.S. Patent No. 3,896,111). Subsequently, it was discovered that certain microbes also produce maytansinoids, such as maytansinol and C-3 maytansinol esters (U.S. Patent No. 4,151,042). Synthetic maytansinol and derivatives and analogues thereof are disclosed, for example, in U.S. Patent Nos. 4,137,230; 4,248,870; 4,256,746; 4,260,608; 4,265,814; 4,294,757; 4,307,016; 4,308,268; 4,308,269; 4,309,428; 4,313,946; 4,315,929; 4,317,821; 4,322,348; 4,331,598; 4,361,650; 4,364,866; 4,424,219; 4,450,254; 4,362,663; and 4,371,533, the disclosures of which are hereby expressly incorporated by reference.

Anti-PUMPCn antibody-maytansinoid conjugates (immunoconjugates)

Anti-PUMPCn antibody-maytansinoid conjugates are prepared by chemically linking an anti-PUMPCn antibody to a maytansinoid molecule without significantly diminishing the biological activity of either the antibody or the maytansinoid molecule. An average of 3-4 maytansinoid molecules conjugated per antibody molecule has shown efficacy in enhancing

cytotoxicity of target cells without negatively affecting the function or solubility of the antibody, although even one molecule of toxin/antibody would be expected to enhance cytotoxicity over the use of naked antibody. Maytansinoids are well known in the art and can be synthesized by known techniques or isolated from natural sources. Suitable maytansinoids are disclosed, for example, in U.S. Patent No. 5,208,020 and in the other patents and nonpatent publications referred to hereinabove. Preferred maytansinoids are maytansinol and maytansinol analogues modified in the aromatic ring or at other positions of the maytansinol molecule, such as various maytansinol esters.

There are many linking groups known in the art for making antibody-maytansinoid conjugates, including, for example, those disclosed in U.S. Patent No. 5,208,020 or EP Patent 0 425 235 B1, and Chari *et al. Cancer Research* 52: 127-131 (1992). The linking groups include disufide groups, thioether groups, acid labile groups, photolabile groups, peptidase labile groups, or esterase labile groups, as disclosed in the above-identified patents, disulfide and thioether groups being preferred.

Conjugates of the antibody and maytansinoid may be made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as toluene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). Particularly preferred coupling agents include N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP) (Carlsson *et al., Biochem. J.* 173:723-737 [1978]) and N-succinimidyl-4-(2-pyridylthio)pentanoate (SPP) to provide for a disulfide linkage.

The linker may be attached to the maytansinoid molecule at various positions, depending on the type of the link. For example, an ester linkage may be formed by reaction with a hydroxyl group using conventional coupling techniques. The reaction may occur at the C-3 position having a hydroxyl group, the C-14 position modified with hyrdoxymethyl, the C-15 position modified with a hydroxyl group, and the C-20 position having a hydroxyl group. In a preferred embodiment, the linkage is formed at the C-3 position of maytansinol or a maytansinol analogue.

Calicheamicin

Another immunoconjugate of interest comprises an anti-PUMPCn antibody conjugated to one or more calicheamicin molecules. The calicheamicin family of antibiotics are capable of producing double-stranded DNA breaks at sub-picomolar concentrations. For the preparation of conjugates of the calicheamicin family, see U.S. patents 5,712,374, 5,714,586, 5,739,116, 5,767,285, 5,770,701, 5,770,710, 5,773,001, 5,877,296 (all to American Cyanamid Company). Structural analogues of calicheamicin which may be used include, but are not limited to, γ_1^I, α_2^I, α_3^I, N-acetyl-γ_1^I, PSAG and θ^I_1 (Hinman *et al. Cancer Research* 53: 3336-3342 (1993), Lode *et al. Cancer Research* 58: 2925-2928 (1998) and the aforementioned U.S. patents to American Cyanamid). Another anti-tumor drug that the antibody can be conjugated is QFA which is an antifolate. Both calicheamicin and QFA have intracellular sites of action and do not readily cross the plasma membrane. Therefore, cellular uptake of these agents through antibody mediated internalization greatly enhances their cytotoxic effects.

Other cytotoxic agents

Other antitumor agents that can be conjugated to the anti-PUMPCn antibodies of the invention include BCNU, streptozoicin, vincristine and 5-fluorouracil, the family of agents known collectively LL-E33288 complex described in U.S. patents 5,053,394, 5,770,710, as well as esperamicins (U.S. patent 5,877,296).

Enzymatically active toxins and fragments thereof which can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin and the tricothecenes. See, for example, WO 93/21232 published October 28, 1993.

The present invention further contemplates an immunoconjugate formed between an antibody and a compound with nucleolytic activity (*e.g.* a ribonuclease or a DNA endonuclease such as a deoxyribonuclease; DNase).

For selective destruction of the tumor, the antibody may comprise a highly radioactive atom. A variety of radioactive isotopes are available for the production of radioconjugated anti-PUMPCn antibodies. Examples include At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32}, Pb^{212} and radioactive isotopes of Lu. When the conjugate is used for

diagnosis, it may comprise a radioactive atom for scintigraphic studies, for example tc^{99m} or I^{123}, or a spin label for nuclear magnetic resonance (NMR) imaging (also known as magnetic resonance imaging, mri), such as iodine-123 again, iodine-131, indium-111, fluorine-19, carbon-13, nitrogen-15, oxygen-17, gadolinium, manganese or iron.

The radio- or other labels may be incorporated in the conjugate in known ways. For example, the peptide may be biosynthesized or may be synthesized by chemical amino acid synthesis using suitable amino acid precursors involving, for example, fluorine-19 in place of hydrogen. Labels such as tc^{99m} or I^{123}, Re^{186}, Re^{188} and In^{111} can be attached via a cysteine residue in the peptide. Yttrium-90 can be attached via a lysine residue. The IODOGEN method (Fraker et al (1978) Biochem. Biophys. Res. Commun. 80: 49-57 can be used to incorporate iodine-123. "Monoclonal Antibodies in Immunoscintigraphy" (Chatal,CRC Press 1989) describes other methods in detail.

Conjugates of the antibody and cytotoxic agent may be made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al. Science* 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026. The linker may be a "cleavable linker" facilitating release of the cytotoxic drug in the cell. For example, an acid-labile linker, peptidase-sensitive linker, photolabile linker, dimethyl linker or disulfide-containing linker (Chari *et al. Cancer Research* 52: 127-131 (1992); U.S. Patent No. 5,208,020) may be used.

Alternatively, a fusion protein comprising the anti-PUMPCn antibody and cytotoxic agent may be made, *e.g.* by recombinant techniques or peptide synthesis. The length of DNA may comprise respective regions encoding the two portions of the conjugate either adjacent one another or separated by a region encoding a linker peptide which does not destroy the desired properties of the conjugate.

In yet another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pre-targeting wherein the antibody-receptor conjugate is

administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.* avidin) which is conjugated to a cytotoxic agent (*e.g.* a radionucleotide).

Antibody Dependent Enzyme Mediated Prodrug Therapy (ADEPT)

The antibodies of the present invention may also be used in ADEPT by conjugating the antibody to a prodrug-activating enzyme which converts a prodrug (*e.g.* a peptidyl chemotherapeutic agent, see WO81/01145) to an active anti-cancer drug. See, for example, WO 88/07378 and U.S. Patent No. 4,975,278.

The enzyme component of the immunoconjugate useful for ADEPT includes any enzyme capable of acting on a prodrug in such a way so as to covert it into its more active, cytotoxic form.

Enzymes that are useful in the method of this invention include, but are not limited to, alkaline phosphatase useful for converting phosphate-containing prodrugs into free drugs; arylsulfatase useful for converting sulfate-containing prodrugs into free drugs; cytosine deaminase useful for converting non-toxic 5-fluorocytosine into the anti-cancer drug, 5-fluorouracil; proteases, such as serratia protease, thermolysin, subtilisin, carboxypeptidases and cathepsins (such as cathepsins B and L), that are useful for converting peptide-containing prodrugs into free drugs; D-alanylcarboxypeptidases, useful for converting prodrugs that contain D-amino acid substituents; carbohydrate-cleaving enzymes such as β-galactosidase and neuraminidase useful for converting glycosylated prodrugs into free drugs; β-lactamase useful for converting drugs derivatized with β-lactams into free drugs; and penicillin amidases, such as penicillin V amidase or penicillin G amidase, useful for converting drugs derivatized at their amine nitrogens with phenoxyacetyl or phenylacetyl groups, respectively, into free drugs. Alternatively, antibodies with enzymatic activity, also known in the art as "abzymes", can be used to convert the prodrugs of the invention into free active drugs (see, *e.g.*, Massey, *Nature* 328: 457-458 (1987)). Antibody-abzyme conjugates can be prepared as described herein for delivery of the abzyme to a tumor cell population.

The enzymes of this invention can be covalently bound to the anti-PUMPCn antibodies by techniques well known in the art such as the use of the heterobifunctional crosslinking reagents discussed above. Alternatively, fusion proteins comprising at least the antigen binding region of an antibody of the invention linked to at least a functionally active portion of an enzyme of the invention can be constructed using recombinant DNA techniques well known in the art (see, *e.g.*, Neuberger *et al.*, *Nature*, 312: 604-608 (1984).

9. Immunoliposomes

The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al* ., J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

10. Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a PRO23203 polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

If the PRO23203 polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993).

The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example,

by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, *supra*.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT™ (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

G. Uses for anti-PRO23203 Antibodies

The anti-PRO23203 antibodies of the invention have various utilities. For example, anti-PRO23203 antibodies may be used in diagnostic assays for PRO23203, *e.g.*, detecting its expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The

antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^{3}H, ^{14}C, ^{32}P, ^{35}S, or ^{125}I, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO23203 antibodies also are useful for the affinity purification of PRO23203 from recombinant cell culture or natural sources. In this process, the antibodies against PRO23203 are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO23203 to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO23203, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO23203 from the antibody.

Potential uses for PRO23203 polypeptides, nucleic acids encoding therefor, and anti-PRO23203 antibodies are in the detection and/or treatment of tumors of prostate tissue origin

Treatment with the Anti-PUMPCn Antibodies

According to the present invention, the anti-PUMPCn antibody that binds PUMPCn on a cell surface is used to treat a PUMPCn-expressing cancer cell such as androgen independent prostate cancer or androgen dependent prostate cancer, and associated metastases. A patient may be diagnosed as having androgen independent prostate cancer in that he no longer responds to anti-androgen therapy and the patient diagnosed as having androgen dependent prostate cancer may be one who responds to anti-androgen therapy. The cancer will generally comprise PUMPCn-expressing cells, such that the anti-PUMPCn antibody is able to bind thereto. While the cancer may be characterized by overexpression of the PUMPCn molecule, the present application further provides a method for treating cancer which is not considered to be an PUMPCn-overexpressing cancer.

Currently, depending on the stage of the cancer, prostate cancer treatment involves one or a combination of the following therapies: surgery to remove the cancerous tissue, radiation therapy, androgen deprivation (e.g., hormonal therapy), and chemotherapy. Anti-

PUMPCn antibody therapy may be especially desirable in elderly patients who do not tolerate the toxicity and side effects of chemotherapy well, in metastatic disease where radiation therapy has limited usefulness, and for the management of prostatic carcinoma that is resistant to androgen deprivation treatment. The tumor targeting and internalizing anti-PUMPCn antibodies of the invention are useful to alleviate PUMPCn-expressing cancers upon initial diagnosis of the disease or during relapse. For therapeutic applications, the anti-PUMPCn antibody can be used alone, or in combination therapy with, e.g., hormones, antiangiogens, or radiolabelled compounds, or with surgery, cryotherapy, and/or radiotherapy, notably for prostate cancers, also particularly where shed cells cannot be reached. Anti-PUMPCn antibody treatment can be administered in conjunction with other forms of conventional therapy, either consecutively with, pre- or post-conventional therapy. Chemotherapeutic drugs such as taxotere® (docetaxel), taxol® (palictaxel), estramustine and mitoxantrone are used in treating metastatic and hormone refractory prostate cancer, in particular, in good risk patients. In the present method of the invention for treating or alleviating cancer, in particular, androgen independent and/or metastatic prostate cancer, the cancer patient can be administered anti-PUMPCn antibody in conjuction with treatment with the one or more of the preceding chemotherapeutic agents. In particular, combination therapy with palictaxel and modified derivatives (see, e.g., EP0600517) is contemplated. The anti-PUMPCn antibody will be administered with a therapeutically effective dose of the chemotherapeutic agent. In another embodiment, the anti-PUMPCn antibody is administered in conjunction with chemotherapy to enhance the activity and efficacy of the chemotherapeutic agent, e.g., paclitaxel. The Physicians' Desk Reference (PDR) discloses dosages of these agents that have been used in treatment of various cancers. The dosing regimen and dosages of these aforementioned chemotherapeutic drugs that are therapeutically effective will depend on the particular cancer being treated, the extent of the disease and other factors familiar to the physician of skill in the art and can be determined by the physician.

In one particular embodiment, an immunoconjugate comprising the anti-PUMPCn antibody conjugated with a cytotoxic agent is administered to the patient. Preferably, the immunoconjugate bound to the PUMPCn protein is internalized by the cell, resulting in increased therapeutic efficacy of the immunoconjugate in killing the cancer cell to which it binds. In a preferred embodiment, the cytotoxic agent targets or interferes with the nucleic acid in the cancer cell. Examples of such cytotoxic agents are described above and include maytansinoids, calicheamicins, ribonucleases and DNA endonucleases.

The anti-PUMPCn antibodies or immunoconjugates are administered to a human patient, in accord with known methods, such as intravenous administration, *e.g.*, as a bolus or by continuous infusion over a period of time, by intramuscular, intraperitoneal, intracerobrospinal, subcutaneous, intra-articular, intrasynovial, intrathecal, oral, topical, or inhalation routes. Intravenous or subcutaneous administration of the antibody is preferred.

Other therapeutic regimens may be combined with the administration of the anti-PUMPCn antibody. The combined administration includes co-administration, using separate formulations or a single pharmaceutical formulation, and consecutive administration in either order, wherein preferably there is a time period while both (or all) active agents simultaneously exert their biological activities. Preferably such combined therapy results in a synergistic therapeutic effect.

It may also be desirable to combine administration of the anti-PUMPCn antibody or antibodies, with administration of an antibody directed against another tumor antigen associated with the particular cancer.

In another embodiment, the antibody therapeutic treatment method of the present invention involves the combined administration of an anti-PUMPCn antibody (or antibodies) and one or more chemotherapeutic agents or growth inhibitory agents, including co-administration of cocktails of different chemotherapeutic agents. Preparation and dosing schedules for such chemotherapeutic agents may be used according to manufacturers' instructions or as determined empirically by the skilled practitioner. Preparation and dosing schedules for such chemotherapy are also described in Chemotherapy Service Ed., M.C. Perry, Williams & Wilkins, Baltimore, MD (1992).

The antibody may be combined with an anti-hormonal compound; *e.g.*, an anti-estrogen compound such as tamoxifen; an anti-progesterone such as onapristone (see, EP 616 812); or an anti-androgen such as flutamide, in dosages known for such molecules. Where the cancer to be treated is androgen independent cancer, the patient may previously have been subjected to anti-androgen therapy and, after the cancer becomes androgen independent, the anti-PUMPCn antibody (and optionally other agents as described herein) may be administered to the patient.

For the prevention or treatment of disease, the dosage and mode of administration will be chosen by the physician according to known criteria. The appropriate dosage of antibody will depend on the type of disease to be treated, as defined above, the severity and course of the disease, whether the antibody is administered for preventive or therapeutic purposes, previous therapy, the patient's clinical history and response to the antibody, and the

discretion of the attending physician. The antibody is suitably administered to the patient at one time or over a series of treatments. Preferably, the antibody is administered by intravenous infusion or by subcutaneous injections. Depending on the type and severity of the disease, about 1 μg/kg to about 50 mg/kg body weight (*e.g.* about 0.1-15mg/kg/dose) of antibody can be an initial candidate dosage for administration to the patient, whether, for example, by one or more separate administrations, or by continuous infusion. A dosing regimen can comprise administering an initial loading dose of about 4 mg/kg, followed by a weekly maintenance dose of about 2 mg/kg of the anti-PUMPCn antibody. However, other dosage regimens may be useful. A typical daily dosage might range from about 1 μg/kg to 100 mg/kg or more, depending on the factors mentioned above. For repeated administrations over several days or longer, depending on the condition, the treatment is sustained until a desired suppression of disease symptoms occurs. The progress of this therapy can be readily monitored by conventional methods and assays and based on criteria known to the physician or other persons of skill in the art.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

The U.S. priority application serial no. 60/235,451 is hereby incorporated by reference in its entirety. All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

EXAMPLE 1

Isolation of cDNA Clones Encoding a Human PRO23203

An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified by GEPIS. Gene expression profiling in silico (GEPIS) is a bioinformatics tool that characterizes genes of interest for new therapeutic targets. GEPIS takes advantage of the vast amount of EST sequence and library information to determine gene expression profiles. GEPIS is based on the assumption that the

expression level of a gene is proportionally correlated with the number of its occurrences in EST databases, and it works by integrating the Incyte EST relational database and Genentech proprietary information in a stringent and statistically meaningful way. In this example, it is used to identify and cross-validate new tumor antigens, although GEPIS can be configured to either perform very specific analyses or broad screening tasks. For the initial screen, GEPIS is used to go from libraries to sequence. The entire Incyte database was used to cluster sequence based on its library information. Breast, colon, lung and prostate were the target organs specified. The sequences found in this initial cluster were then subjected to a screen for secreted and transmembrane containing domains. The remaining sequences were then screened for novelty and those individual sequences identified. In a final step, each individual sequence was then put through a GEPIS screen, this time going from sequence to library, confirming its expression profile in the original target tissue (Figure 3). Using this type of screening bioinformatics, DNA182753 was identified, and PCR primers designed using this sequence were used to screen libraries for the full length clone.

RNA for construction of cDNA libraries was then isolated from human prostate tissue. The cDNA libraries used to isolate the cDNA clones encoding human PRO23203 were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI.

Oligonucleotides probes based upon the above described EST sequence were then synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO23203. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, *supra*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The oligonucleotide probes employed were as follows:

forward PCR primer 5'-GATATTTGTTTCTCAACATGGCTTATCAGCAGG-3' (SEQ ID

NO:3)

reverse PCR primer 5'-TCTCTGACCTTCTCATCGGTAAGCAGAGG-3' (SEQ ID NO:4)

hybridization probe

5'-TCTTTTGCAGCTTTGCAGATACCCAGACTGAGCTGGAACTGGA-3' (SEQ ID NO:5)

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 188-190 and a stop signal at nucleotide positions 1550-1552 (Figure 1, SEQ ID NO:1). The predicted polypeptide precursor is 454 amino acids long, has a calculated molecular weight of approximately 52008 daltons and an estimated pI of approximately 8.83. Analysis of the full-length PRO23203 sequence shown in Figure 2 (SEQ ID NO:2) evidences the presence of a variety of important polypeptide domains, wherein the locations given for those important polypeptide domains as shown immediately below, are approximate as described above.

Signal peptide: None

Transmembrane domain:

210-230
256-278
302-321
360-382
391-412
430-450

N-glycosylation site: 256-259

cAMP- and cGMP-dependent protein kinase phosphorylation site: 29-32

Tyrosine kinase phosphorylation site: 416-424

N-myristoylation site.
8-13
24-29
34-39
193-198
274-279

The ECDs and ICDs likely lie outside of the amino acids delineating the predicted TMs above.

Clone DNA185171-2994 has been deposited with ATCC on September 26, 2000 and is assigned ATCC deposit no. PTA-2513.

An analysis of the protein database (version 35.45 SwissProt 35), using the ALIGN-2 sequence alignment analysis of the full-length sequence shown in Figure 2 (SEQ ID NO:2),

evidenced sequence identity between the PRO23203 amino acid sequence and the following sequences: AK001691_1.

EXAMPLE 2

Use of PRO23203 as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO23203 as a hybridization probe.

DNA comprising the coding sequence of full-length or mature PRO23203 is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO23203) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO23203-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO23203 can then be identified using standard techniques known in the art.

EXAMPLE 3

Expression of PRO23203 in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO23203 by recombinant expression in *E. coli*.

The DNA sequence encoding PRO23203 is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO23203 coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the

methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO23203 protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO23203 may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO23203 is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30EC with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g $(NH_4)_2SO_4$, 0.71 g sodium citrate•2H2O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM $MgSO_4$) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column

equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO23203 polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

EXAMPLE 4

Expression of PRO23203 in mammalian cells

This example illustrates preparation of a potentially glycosylated form of PRO23203 by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO23203 DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO23203 DNA using ligation methods such as

described in Sambrook et al., supra. The resulting vector is called pRK5-PRO23203.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μg pRK5-PRO23203 DNA is mixed with about 1 μg DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 μl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M $CaCl_2$. To this mixture is added, dropwise, 500 μl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM $NaPO_4$, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μCi/ml ^{35}S-cysteine and 200 μCi/ml ^{35}S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO23203 polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO23203 may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μg pRK5-PRO23203 DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μg/ml bovine insulin and 0.1 μg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO23203 can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO23203 can be expressed in CHO cells. The pRK5-PRO23203 can be transfected into CHO cells using known reagents such as $CaPO_4$ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S-methionine. After

determining the presence of PRO23203 polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO23203 can then be concentrated and purified by any selected method.

Epitope-tagged PRO23203 may also be expressed in host CHO cells. The PRO23203 may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO23203 insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO23203 can then be concentrated and purified by any selected method, such as by Ni^{2+}-chelate affinity chromatography.

PRO23203 may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Quiagen), Dosper® or Fugene® (Boehringer Mannheim). The cells are grown as described in Lucas et al., supra. Approximately 3 x 10^7 cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into water bath and

mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 μm filtered PS20 with 5% 0.2 μm diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with 3 x 10^5 cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2 x 10^6 cells/mL. On day 0, the cell number and pH are determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μm filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 μL of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman

degradation.

EXAMPLE 5

Expression of PRO23203 in Yeast

The following method describes recombinant expression of PRO23203 in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO23203 from the ADH2/GAPDH promoter. DNA encoding PRO23203 and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO23203. For secretion, DNA encoding PRO23203 can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO23203 signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO23203.

Yeast cells, such as yeast strain ANTIBODY110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO23203 can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO23203 may further be purified using selected column chromatography resins.

EXAMPLE 6

Expression of PRO23203 in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO23203 in Baculovirus-infected insect cells.

The sequence coding for PRO23203 is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO23203 or the desired portion of the coding sequence of PRO23203 such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5'

primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO23203 can then be purified, for example, by Ni^{2+}-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM $MgCl_2$; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni^{2+}-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni^{2+}-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His_{10}-tagged PRO23203 are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO23203 can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

EXAMPLE 7

Preparation of Antibodies that Bind PRO23203

This example illustrates preparation of monoclonal antibodies which can specifically bind PRO23203.

Techniques for producing the monoclonal antibodies are known in the art and are

described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO23203, fusion proteins containing PRO23203, and cells expressing recombinant PRO23203 on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO23203 immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO23203 antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO23203. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO23203. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO23203 is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO23203 monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 8

Purification of PRO23203 Polypeptides Using Specific Antibodies

Native or recombinant PRO23203 polypeptides may be purified by a variety of

standard techniques in the art of protein purification. For example, pro-PRO23203 polypeptide, mature PRO23203 polypeptide, or pre-PRO23203 polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO23203 polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO23203 polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO23203 polypeptide by preparing a fraction from cells containing PRO23203 polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO23203 polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO23203 polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO23203 polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO23203 polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO23203 polypeptide is collected.

EXAMPLE 9

Drug Screening

This invention is particularly useful for screening compounds by using PRO23203 polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO23203 polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with

recombinant nucleic acids expressing the PRO23203 polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO23203 polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO23203 polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO23203 polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO23203 polypeptide or fragment thereof and assaying (I) for the presence of a complex between the agent and the PRO23203 polypeptide or fragment, or (ii) for the presence of a complex between the PRO23203 polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO23203 polypeptide or fragment is typically labeled. After suitable incubation, free PRO23203 polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO23203 polypeptide or to interfere with the PRO23203 polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO23203 polypeptide, the peptide test compounds are reacted with PRO23203 polypeptide and washed. Bound PRO23203 polypeptide is detected by methods well known in the art. Purified PRO23203 polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO23203 polypeptide specifically compete with a test compound for binding to PRO23203 polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO23203 polypeptide.

EXAMPLE 10

Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active

polypeptide of interest (*i.e.*, a PRO23203 polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO23203 polypeptide or which enhance or interfere with the function of the PRO23203 polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO23203 polypeptide, or of an PRO23203 polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO23203 polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO23203 polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO23203 polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO23203 polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO23203 polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

EXAMPLE 11

Expression of PUMPCn by Taqman™ Analysis

As an intitial study, expression of PUMPCn in human cDNA libraries [Figure 4] prepared from normal as well as tumor tissues and cell lines was analyzed by Taqman™. Fifty

nanograms of each cDNA library were used. The following primers to the 3' end of PUMPCn were used in the Taqman analysis.

Forward Primer (19 mer): GCC AGC CGG CAG GTT TAT A (SEQ ID NO. 6)

Reverse primer (19 mer): ATT CAA CTG GCG GGC AAG T (SEQ ID NO. 7)

Probe (26 mer): TGC AGC AAC AAT ATT CAA GCG CGA CA (SEQ ID NO. 8)

The TaqMan™ reaction is a fluorescent PCR-based technique which makes use of the 5' exonuclease activity of Taq DNA polymerase enzyme tomonitor amplification in real time. Two oligonucleotide primers are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the amplification reaction, the Taq DNA polymerase enzyme cleaves the probe in a template-dependent manner. The resultant probe fragments disassociate in solution, and signal from the released reporter dye is free from the quenching effect of the second fluorophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data. The results of the TaqManTM reaction are reported in delta (Δ) Ct units. TaqMan™ assay data are initially expressed as Ct, or the threshold cycle. This is defined as the cycle at which the reporter signal accumulates above the background level of fluorescence. The ΔCt values are used as quantitative measurement of the relative number of starting copies of a particular target sequence in a nucleic acid sample when comparing cancer results to normal human results. One unit corresponds to 1 PCR cycle or approximately a 2-fold amplification relative to normal, two units corresponds to 4-fold, 3 units to 8-fold amplification and so on.

The results of the Taqman™ analysis which were normalized against expression of β-actin, are shown in Figure 4 and Figure 5. In Figure 4, the PUMPCn expression in other tissues is represented relative to the expression in normal prostate (prostate given a value of 1.0). In Figure 4, PUMPCn is highly expressed in the prostate with expression also detected in colon as well as in lung tumors and liver tumors. Figure 5 shows fold expression of PUMPCn in tumor versus normal tissue in a subset of the non-prostate cDNA libraries shown in Figure 4. PUMPCn was found to be overexpressed in lung, esophagus and liver tumors.

As a follow-up to the Taqman analysis, *in situ* hybridization was performed on various normal and cancerous tissues to provide a more defined look at expression. The expression

pattern seen here was consistent with the results from the *in situ* hybridization analysis described in Example 12 below.

EXAMPLE 12

Expression of PUMPCn by *In situ* Hybridization

In situ hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression, analyze the tissue distribution of transcription, identify and localize viral infection, follow changes in specific mRNA synthesis and aid in chromosome mapping.

In situ hybridization was performed following an optimized version of the protocol by Lu and Gillett, *Cell Vision* 1: 169-176 (1994), using PCR generated α-^{33}P riboprobes. Briefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinated in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. An α-^{33}P UTP antisense riboprobe was generated from a PCR product designed to have T3 and T7 RNA polymerase promoters at either end and hybridized to the tissues at 55°C overnight. The slides were dipped in Kodak NTB2 nuclear track emulsion and exposed for 4 weeks.

Riboprobe synthesis

6.0 µl (125 mCi) of (Amersham BF 1002, SA<2000 Ci/mmol) were speed vac dried. To each tube containing dried α-^{33}P UTP, the following ingredients were added: 2.0 ml 5x transcription buffer; 1.0 µl DTT (100mM); 2.0 µl NTP mix (2.5 mM : 10 µl; each of 10 mM GTP, CTP & ATP + 10 µl H_2O); 1.0 µl UTP (50 µM); 1 µl Rnasin; 1.0 µl DNA template (1 µg); 1.0 µl H_2O.

The tubes were incubated at 37°C for one hour. 1.0 µl RQ1 DNase were added, followed by incubation at 37°C for 15 minutes. 90 µl TE (10 mM Tris pH 7.6/1mM EDTA pH 8.0) were added, and the mixture was pipetted onto DE81 paper. The remaining solution was loaded in a Microcon-50 ultrafiltration unit, and spun in a Heraeus Sepatech Centrifuge 28RS at 12,000 RPM (6 minutes). The filtration unit was inverted over a second tube and spun at 3500 RPM (3 minutes). After the final recovery spin, 100 µl TE were added. 1 µl of the final product was pipetted on DE81 paper and counted in 6 ml of Biofluor II on a Beckman LS 5000 TD scintillation counter.

The probe was run on a TBE/urea gel. 1-3 µl of the probe or 5 µl of RNA Mrk III were added to 3 µl of loading buffer. After heating on a 95°C heat block for three minutes,

the probe was immediately placed on ice. The wells of gel were flushed, the sample loaded, and run at 180-250 volts for 45 minutes. The gel was wrapped in saran wrap and exposed to XAR film with an intensifying screen in -70°C freezer one hour to overnight.

Hybridization

Pretreatment of frozen sections. The slides were removed from the freezer, placed on aluminum trays and thawed at room temperature for 5 minutes. The trays were placed in 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H_2O). After deproteination in 0.5 µg/ml proteinase K for 10 minutes at 37°C (12.5 µl of 10 mg/ml stock in 250 ml prewarmed RNase-free RNAse buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, 100% ethanol, 2 minutes each.

Pretreatment of paraffin-embedded sections. The slides were deparaffinized through three changes of xylenen, 100% ethanol and rehydrated through graded ethanols to water, placed in SQ H_2O and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinated in 20 µg/ml proteinase K (500 µl of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) - human embryo, or 8 x proteinase K (100 µl in 250 ml Rnase buffer, 37°C, 30 minutes) - formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

Prehybridization. The slides were laid out in plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper. The tissue was covered with 50 µl of hybridization buffer (10% Dextran sulfate, 50% formamide, 2X SSC) and incubated at 42°C for 1-4 hours.

Hybridization. 1.0 X 10^6 cpm probe and 1.0 µl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 µl hybridization buffer were added to the probe/tRNA mix per slide. After vortexing, 50 µl ^{33}P mix were added to 50 µl prehybridization on slide. The slides were incubated overnight at 55°C.

Washes. Washing was done 2x10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25M EDTA, V_f= 4L), followed by RNaseA treatment at 37°C for 30 minutes (500 µl of 10 mg/ml in 250 ml Rnase buffer - 20 µg/ml). The slides were washed 2x10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55EC, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f= 4L).

Human and chimp studies were performed using a probe against the 3' segment of PUMPCn. PCR primers were designed to amplify portions of the gene which can serve as templates for in vitro transcription of radioactively-labeled (^{33}P) single-stranded complementary riboprobes. Sense (control) riboprobes are generated by transcription using T7 RNA polymerase, which recognizes a 27 nucleotide sequence for the T7 promoter appended onto the 5' end of the upper PCR primer. Antisense (experimental) riboprobes are generated by transcription using T3 RNA polymerase, which recognizes a 27 nucleotide sequence for the T3 promoter appended onto the 5' end of the lower PCR primer. The PUMPCn probe used in the first study [ISH2000-121] was designed to span nucleotides 1149-1503 of DNA 185171-2994 [SEQ ID NO. 1]

Upper primer (SEQ ID NO. 9):

5' GGATTCTAAT ACGACTCACT ATAGGGCCTG CTTACCGATC AGAAGGTC 3'

lower primer (SEQ ID NO. 10):

5' CTATGAAATT AACCCTCACT AAAGGGAGGC AAAACAAGAG CAAGAACA 3'

The probe sequence is:

5'

CTGCTTACCGATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTATCAG
CAGGTTCATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTG
AAATGTATATCTCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTC
ACTTCTATCCCTTCAGTGAGCAATGCTTTAAACTGGAGAGAATTCAGTTTTATTCA
GTCTACACTTGGATATGTCGCTCTGCTCATAAGTACTTTCCATGTTTTAATTTATGG
ATGGAAACGAGCTTTTGAGGAAGAGTACTACAGATTTTATACACCACCAAACTTT
GTTCTTGCTCTTGTTTTGCC -3' (SEQ ID NO. 11)

The human studies were repeated using a probe against the 5' part of PUMPCn. This probe gave the same results as the probe to the 3' part of the gene. The probe used in this second study was designed to span nucleotides 14-566 of DNA 185171 (SEQ ID NO. 1)

upper primer (SEQ ID NO. 12):

5' GGATTCTAATACGACTCACTATAGGGCATGGAACAGTATATGGAAAGC 3'

lower primer(SEQ ID NO. 13):

5' CTATGAAATTAACCCTCACTAAAGGGACTG GGTACTGGTTTATCCTC

The probe sequence is: 5'
ATGGAACAGTATATGGAAAGCTCCCAAGAAAGTGAAGAGAGGAAATTGGAAAA
TTGTGAGTGGACCTTCTGATACTGCTCCTCCTTGCGTGGAAAAGGGGAAAGAAC
TGCATGCATATTATTCAGCGTCCTATATTCAAAGGATATTCTTGGTGATCTTGGA
AGTGTCCGTATCATGGAATCAATCTCTATGATGGGAAGCCCTAAGAGCCTTAGT
GAAACTTGTTTACCTAATGGCATAAATGGTATCAAAGATGCAAGGAAGGTCACT
GTAGGTGTGATTGGAAGTGGAGATTTTGCCAAATCCTTGACCATTCGACTTATTA
GATGCGGCTATCATGTGGTCATAGGAAGTAGAAATCCTAAGTTTGCTTCTGAATT
TTTTCCTCATGTGGTAGATGTCACTCATCATGAAGATGCTCTCACAAAAACAAAT
ATAATATTTGTTGCTATACACAGAGAACATTATACCTCCCTGTGGGACCTGAGAC
ATCTGCTTGTGGGTAAAATCCTGATTGATGTGAGCAATAACATGAGGATAAACC
AGTACCCAG 3' (SEQ ID NO. 14)

Results

There is high expression of DNA185171-2994 in a tissue panel containing a series of benign prostate, primary prostate carcinoma, and metastatic prostate carcinoma; and in an array of tumors from a variety of locations. Hybridization signal is seen in tumors of lung, colon, bladder, prostate, and endometrium. In an array of normal tissues from a variety of locations, hybridization signal is seen in prostate and ovary. Note that all tissues were hybridized at the same time; and signal in the prostate tissues is consistent and often comparatively very high, although intensities are variable.

Table 4 summarizes PUMPCn expression as relative signal intensity in a tissue panel. There was high expression of DNA185171-2994 in a tissue panel containing a series of benign prostate, primary prostate carcinoma, and metastatic prostate carcinoma. Hybridization signal was seen in normal prostate. In an array of prostate tissue, the highest signal intensities were often associated with metastatic tumors. This association was not noticed in studies with other molecules believed to be involved in prostate tumors. We quantitated DNA185171-2994 and a control molecule previously shown to be highly expressed in prostate tissue only. Expression levels were analyzed by phosphorimager analysis of the ^{33}P-labeled hybridization probes. The DNA185171-2994 signal from each element of the tissue array was divided by the control signal from the same element. The DNA185171-2994/control ratio was 3 times higher in the metastatic tumor group compared to the primary carcinoma group. Thus, DNA185171-2994 may be an especially attractive target for metastatic tumors.

TABLE 4

PUMPCn Expression in Human Tissues and Tumors - In Situ Hybridization

Site, tissue, or tumor type	Signal Intensity (Number of cases)			
	-	-/+	+	++
Adrenal	2			
Aorta	1			
Brain (neurons in cortex, cerebellum, brainstem)	2		2	
Breast (normal/benign)	15			
Breast, invasive carcinoma	14		1	
Colon (normal/benign)	6	2		
Colon adenocarcinoma	3	2	4	
Eye retina	2			
Gallbladder (epithelium)	1	1	1	
Heart	2			
Kidney (adult normal/benign)	4			
Kidney, embryonic			1	
Kidney, renal cell carcinoma	3			
Liver (normal/benign)	3			
Liver hepatocellular carcinoma	2			
Lung (adult normal/benign)	23	1		
Lung fetal	1			
Lung adenocarcinoma	4		7	5
Lung carcinoid tumor			1	
Lung non-small cell carcinoma, not otherwise specified	5	1	5	4
Lung squamous cell carcinoma		2	3	
Lymph node (normal/benign)	1			
Lymphoma	2			
Ovary (normal/benign)				
Ovary adenocarcinoma	1	1		
Pancreas (normal/benign)				
Pancreas adenocarcinoma	1		2	
Placenta	1			
Prostate (normal, benign)	5	4	15	19
Prostate PIN	1	1	1	9
Prostate primary adenocarcinoma	4	8	12	86
Prostate metastatic adenocarcinoma	12	4	3	26
Prostate transitional cell carcinoma			1	
Skin (normal/benign)	2			
Skin, malignant melanoma	3			
Small intestine, adult	3			
Small intestine, embryonic		1		
Spleen	4			
Stomach (normal/benign)	2			
Stomach adenocarcinoma	1			
Thymus				
Thyroid (epithelium)	1	1	1	
Testis	1			
Tonsil	2			
Urinary bladder (urothelium)	4		1	
Uterus (normal/benign)				

Uterus-endometrial adenocarcinoma		1	1	1
LnCAP human prostate cancer cells (cell pellet, xenograft)				2
SKBr3 human breast cancer cells	1			
MDA231 human breast cancer cells	1			
Inflammatory stromal cells (probably histiocytes)				
			1	
		1	1	2
			3	
			2	1
			1	
			1	

EXAMPLE 13

Monoclonal Antibodies to PUMPCn

Cell lines and transfections– 293 cells are a human immortalized embryonic kidney cell line (ATCC reference CRL1573, PC-3 is a human prostate cancer cell line (ATCC reference CRL1435) and SV-T2 is a mouse embryonic fibroblast cell line. Growth conditions were according to ATCC guidelines. For all cell lines, 1μg DNA or 3μg DNA were transfected into 6-well or 10cm dishes, respectively, using Effectene (Qiagen). Cells were harvested or fixed 48 hours post transfection.

cDNA constructs - Full length PUMPCn (DNA #185171) containing amino acids 1-454 was subcloned as a blunted XbaI/PstI fragment from SV40 #185171 into the EcoRV site of a modified version of pCDNA3 (Invitrogen) containing a myc epitope. Myc PUMPCn 5' containing amino acids 1-194 and myc PUMPCn 3' containing amino acids 193-454 were derived from pCDNA3 myc full length PUMPCn digested with BamHI and subcloning into BamHI digested pCDNA3 myc or recircularizing original vector, respectively. The construct full length gD PUMPCn codes for amino acids 1-454 with a gD epitope tag at its amino terminus and was derived by subcloning a SalI/XbaI fragment from pCDNA3 myc full length PUMPCn (removing the myc tag) into gD vec 806 digested with XhoI/XbaI. Untagged full length PUMPCn under a CMV promoter was constructed by subcloning a SalI/NotI fragment from pCDNA3 myc full length PUMPCn (removing the myc tag) into the XhoI/NotI sites of pCDNA3.1-(Invitrogen). All constructs were confirmed by DNA sequencing using an ABI sequencer.

Antibodies and immunological procedures - Female Balb/c mice were immunized with an N-terminal fragment corresponding to amino acids 25-213 starting with met, lys and ending

with trp arg [see Figure 2; SEQ ID NO. 2]. The mice were injected i.v. in the footpads. Monoclonal Abs were screened by ELISA against the immunizing peptide. Clones which were positive in the ELISA were further tested by Western blot analysis and immunohistochemistry (IHC).

Fresh-frozen sections and formalin-fixed, paraffin-embedded sections of LnCAP human prostate carcinoma cells were examined by IHC. In situ hybridization studies done previously showed that LnCAP cells showed high levels of PUMPCn mRNA, similar to the levels detected in tissue sections of prostate adenocarcinoma specimens. The IHC was performed by the avidin-biotin/peroxidase complex method using the anti-PUMPCn monoclonal antibodies as primary antibody.

Monoclonal antibodies to the myc epitope were purchased from Invitrogen; the monoclonal antibody to the gD epitope is 5B6 mAb (Genentech). For western analysis, approximately 50μg of total soluble protein or total cell lysate was resolved by SDS-gel electrophoresis performed using precasted gels according to the manufacturer's instructions (Novex) and blotted to PVDF membrane. Soluble protein lysates were prepared by lysing transfected cells in 5 volumes of Triton X-100 buffer [20mM tris-HCl, pH 8.0, 1% Triton X-100, 137mM NaCl, 10% glycerol, 1mM EGTA, 1.5mM MgCl2, 1mM dithiothreitol (DTT), 1mM sodium vanadate, 50mM NaF, 1mM Pefabloc, 10μg/ml each of Aprotinin, pepstatin and leupeptin] and clarifying by centrifugation at 14,000 rpm for 10 minutes. Total cell lysates were prepared by lysing directly in 1X SDS sample buffer. For western blotting, antibodies were used at a final concentration of 2μg/ml and developed using the ECL system (Amersham). Immunofluorescent detection of tagged PUMPCn in fixed whole cells was carried out on transfected 293 or PC-3 cells grown on coverslips, fixed in 4% paraformaldehyde and stained using mouse monoclonal antibodies to the respective epitope tags. The cells were visualized with either FITC or Texas Red labeled secondary antibodies (Jackson Labs) and co-stained with DAPI for nuclei detection.

293 and PC-3 cells were plated onto coverslips in 6-well dishes and transfected with myc tagged full length PUMPCn, myc PUMPCn 5', myc PUMPCn 3' and gD PUMPCn using Effectene. At 24 hours post transfection, media was changed and incubated for an additional 24 hours. Cells were fixed using 4% paraformaldehyde at room temperature for 10 minutes, washed 4 times in PBS and stored in foil at 4°C till processing. Coverslips were stained for anti-myc or anti-gD reactivity and visualized using FITC – coupled donkey anti-mouse antibody.

293, PC-3 and SV-T2 cells were transfected with myc tagged full length PUMPCn,

myc PUMPCn 5', myc PUMPCn 3' and gD PUMPCn using Effectene. At 24 hours post transfection, media was changed and incubated for an additional 24 hours. Transfected cells were lysed as described and analyzed for expression of epitope tagged PUMPCn.

The monoclonal antibodies generated against the N-terminal fragment of PUMPCn (amino acids 25-213) were first tested for binding to recombinant PUMPCn used as the immunogen and then to endogenous PUMPCn present in normalized cell lysates prepared from the following cell lines: 293, SW480, Colo205, HPAC, SW780, LnCAP, and HPAFII.

Results

The monoclonal antibodies generated to the N-terminal fragment amino acids 25-213 are as follows:

mAb	Clone	Isotype
3248	2H6.2.1	IgG1, K
3249	3A9.2.1	IgG2b, K
3250	3B4.2.1	IgG2b, K
3251	3H9.1.1	IgG2b, K
3252	5D11.1.1	IgG2a, K
3253	5F12.1.1	IgG2b, K
3254	5G9.1.1	IgG2a, K
3255	6B4.1.1	IgG2b, K
3256	7B6.2.1	IgG1, K

The clones (hybridoma cell lines) tested by IHC were 3248 (2H6.2.1), 3249 (3A9.2.1), 3250 (3B4.2.1), 3251 (3H9.1.1), 3252 (5D11.1.1), 3253 (5F12.1.1), 3254 (5G9.1.1), 3255 (6B4.1.1), 3256 (7B6.2.1). The IHC demonstrated moderate to strong immunoreactivity associated with LnCAP cell plasma membranes in fresh frozen sections using MAb 3248, 3249, 3251, 3253 and 3255. In the formalin-fixed, paraffin embedded LnCAP cell sections, MAbs 3248 and 3249 produced positive immunoreactivities which were specifically localized to the plasma membranes. This membrane localization is consistent with the predicted transmembrane structure based on sequence analysis and supports the validity of PUMPCn as an accessible therapeutic target. The monoclonal antibodies described here can be used to demonstrate the localization and expression level of PUMPCn in tissue sections.

From the results in the Examples above, antibodies and hybridization probes to DNA185171-2994 are useful for detecting DNA185171-2994 in tumors that express PUMPCn. Antibodies, naked or conjugated to cytotoxic agents are useful for the treatment of cancers wherein the tumors express PUMPCn on the cell surface, in particular, prostate cancer. In another instance, antibodies and hybridization probes to DNA185171-2994 may be useful in detecting if a metastatic tumor has origins in prostate tissue. In a further instance, antibodies to DNA185171-2994 may be useful in the treatment of tumors of prostate origin that have since metastasized to other areas of the body.

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC):

Material	ATCC Accession No.	Deposit Date
DNA185171-2994	PTA-2513	September 26, 2000

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposit will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the

construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule which comprises DNA having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO23203 polypeptide comprising the sequence of amino acid residues from about 1 to about 454 of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a).

2. The isolated nucleic acid molecule of Claim 1 comprising the sequence of nucleotide positions from about 188 to about 1549 of Figure 1 (SEQ ID NO:1).

3. The isolated nucleic acid molecule of Claim 1 comprising the nucleotide sequence of Figure 1 (SEQ ID NO:1).

4. The isolated nucleic acid molecule of Claim 1 comprising a nucleotide sequence that encodes the sequence of amino acid residues from about 1 to about 454 of Figure 2 (SEQ ID NO:2).

5. An isolated nucleic acid molecule comprising DNA which comprises at least about 80% sequence identity to (a) a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994), or (b) the complement of the DNA molecule of (a).

6. The isolated nucleic acid molecule of Claim 5 comprising DNA encoding the same mature polypeptide encoded by the human protein cDNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994).

7. An isolated nucleic acid molecule comprising DNA which comprises at least about 80% sequence identity to (a) the full-length polypeptide coding sequence of the human protein cDNA deposited with the ATCC on September 26, 2000 under ATCC Deposit No. PTA-2513 (DNA185171-2994), or (b) the complement of the coding sequence of (a).

8. The isolated nucleic acid molecule of Claim 7 comprising the full-length polypeptide coding sequence of the human protein cDNA deposited with the ATCC on

September 26, 2000 under ATCC Deposit No.PTA-2513 (DNA185171-2994).

9. An isolated nucleic acid molecule encoding a PRO23203 polypeptide comprising DNA that hybridizes to the complement of the nucleic acid sequence that encodes amino acids 1 to about 454 of Figure 2 (SEQ ID NO:2).

10. The isolated nucleic acid molecule of Claim 9, wherein the nucleic acid that encodes amino acids 1 to about 454 of Figure 2 (SEQ ID NO:2) comprises nucleotides 188 to about 1549 of Figure 1 (SEQ ID NO:1).

11. The isolated nucleic acid molecule of Claim 9, wherein the hybridization occurs under stringent hybridization and wash conditions.

12. An isolated nucleic acid molecule comprising at least about 1204 nucleotides and which is produced by hybridizing a test DNA molecule under stringent hybridization conditions with (a) a DNA molecule which encodes a PRO23203 polypeptide comprising a sequence of amino acid residues from 1 to about 454 of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a), and isolating the test DNA molecule.

13. The isolated nucleic acid molecule of Claim 12, which has at least about 80% sequence identity to (a) or (b).

14. A vector comprising the nucleic acid molecule of Claim 1.

15. The vector of Claim 14, wherein said nucleic acid molecule is operably linked to control sequences recognized by a host cell transformed with the vector.

16. A nucleic acid molecule deposited with the ATCC under accession number PTA-2513 (DNA185171-2994).

17. A host cell comprising the vector of Claim 14.

18. The host cell of Claim 17, wherein said cell is a CHO cell.

19. The host cell of Claim 17, wherein said cell is an *E. coli*.

20. The host cell of Claim 17, wherein said cell is a yeast cell.

21. A process for producing a PRO23203 polypeptide comprising culturing the host cell of Claim 17 under conditions suitable for expression of said PRO23203 polypeptide and recovering said PRO23203 polypeptide from the cell culture.

22. An isolated PRO23203 polypeptide comprising an amino acid sequence comprising at least about 80% sequence identity to the sequence of amino acid residues from about 1 to about 454 of Figure 2 (SEQ ID NO:2).

23. The isolated PRO23203 polypeptide of Claim 22 comprising amino acid residues 1 to about 454 of Figure 2 (SEQ ID NO:2).

24. An isolated PRO23203 polypeptide having at least about 80% sequence identity to the polypeptide encoded by the cDNA insert of the vector deposited with the ATCC on September 26, 2000 as ATCC Deposit No. PTA-2513 (DNA185171-2994).

25. The isolated PRO23203 polypeptide of Claim 24 which is encoded by the cDNA insert of the vector deposited with the ATCC on September 26, 2000 as ATCC Deposit No. PTA-2513 (DNA185171-2994).

26. An isolated PRO23203 polypeptide comprising the sequence of amino acid residues from 1 to about 454 of Figure 2 (SEQ ID NO:2), or a fragment thereof sufficient to provide a binding site for an anti-PRO23203 antibody.

27. An isolated polypeptide produced by (i) hybridizing a test DNA molecule under stringent conditions with (a) a DNA molecule encoding a PRO23203 polypeptide comprising the sequence of amino acid residues from 1 to about 454 of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a), (ii) culturing a host cell comprising said test DNA molecule under conditions suitable for the expression of said polypeptide, and (iii) recovering said polypeptide from the cell culture.

28. The isolated polypeptide of Claim 27, wherein said test DNA has at least about 80% sequence identity to (a) or (b).

29. A chimeric molecule comprising a PRO23203 polypeptide fused to a heterologous amino acid sequence.

30. The chimeric molecule of Claim 29, wherein said heterologous amino acid sequence is an epitope tag sequence.

31. The chimeric molecule of Claim 29, wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

32. An antibody which specifically binds to a PRO23203 polypeptide.

33. The antibody of Claim 32, wherein said antibody is a monoclonal antibody.

34. The antibody of Claim 32, wherein said antibody is a humanized antibody.

35. The antibody of Claim 32, wherein said antibody is an antibody fragment.

36. The antibody of Claim 32 wherein said antibody is linked to a cytotoxic agent.

37. The antibody of Claim 32 wherein said antibody is linked to a label.

38. An agonist to a PRO23203 polypeptide.

39. An antagonist to a PRO23203 polypeptide.

40. A composition of matter comprising (a) a PRO23203 polypeptide, (b) an agonist to a PRO23203 polypeptide, (c) an antagonist to a PRO23203 polypeptide, or (d) an anti-PRO23203 antibody in admixture with a pharmaceutically acceptable carrier.

41. An isolated nucleic acid molecule which comprises a nucleotide sequence having at least about 80% sequence identity to (a) a DNA molecule encoding amino acids 1 to

X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a).

42. The isolated nucleic acid of Claim 39 which comprises (a) a nucleotide sequence encoding amino acids 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2), or (b) the complement of the nucleotide sequence of (a).

43. An isolated nucleic acid molecule which comprises (a) a nucleotide sequence encoding a polypeptide scoring at least about 80% positives when compared with an amino acid sequence of residues from about 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2), or (b) the complement of the nucleotide sequence of (a).

44. An isolated soluble PRO23203 polypeptide comprising an amino acid sequence having at least about 80% sequence identity to amino acids 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2).

45. The isolated soluble PRO23203 polypeptide of Claim 42 which comprises amino acids 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2).

46. An isolated soluble PRO23203 polypeptide comprising an amino acid sequence which scores at least about 80% positives when compared with the amino acid sequence of amino acids 1 to X of Figure 2 (SEQ ID NO:2), where X is any amino acid from 205 to 214 of Figure 2 (SEQ ID NO:2).

47. A method of treating prostate cancer in a mammal comprising administering to the mammal a therapeutically effective amount of an anti-PRO23203 antibody.

48. The method of Claim 47, wherein the prostate cancer is androgen independent prostate cancer.

49. The method of Claim 47, wherein the cancer is of prostate origin that has metastasized to another part of the body.

50. The method of Claim 47, wherein the antibody is an antibody fragment.

51. The method of Claim 47, wherein the antibody fragment is a Fab fragment.

52. The method of Claim 47, wherein the antibody is not conjugated with a cytotoxic agent.

53. The method of Claim 50, wherein the antibody fragment is not conjugated with a cytotoxic agent.

54. The method of Claim 47, wherein the antibody is conjugated with a cytotoxic agent.

55. The method of Claim 47, further comprising administering to said mammal a chemotherapeutic agent.

56. An article of manufacture comprising a container and a composition contained therein, wherein the composition comprises an anti-PRO23203 antibody, and further comprising a package insert indicating that the composition can be used to treat prostate cancer.

57. The article of manufacture of Claim 56 wherein the prostate cancer is androgen independent prostate cancer

58. The article of manufacture of Claim 56 wherein the package insert further indicates treating the patient with a chemotherapeutic agent.

59. A method of diagnosing the presence of prostate cancer in a mammal, said method comprising contacting a tissue of said mammal with an anti-PRO23203 antibody and detecting the binding of said antibody to a component of said tissue, wherein the binding of said antibody to a component of said tissue is indicative of the presence of prostate cancer in

said mammal.

60. The method of Claim 59, wherein said contacting is performed ex vivo.

61. An article of manufacture comprising a container and a composition contained therein, wherein said composition comprises an anti-PRO23203 antibody and further comprising a package insert indicating that the composition can be used in the diagnosis of prostate cancer.

62. A method of diagnosing the presence of prostate cancer in a mammal, said method comprising contacting a microarray diagnostic with a DNA185171-2994 probe, detecting and quantifying hybridization of said DNA185171-2994 probe in prostate cancer tissue compared with normal tissue and determining if DNA185171-2994 is overexpressed.

TTCTGCTATAGAGATGGAACAGTATATGGAAAGCTCCCAAGAAAGTGAAGAGAGGAAATT
GGAAAATTGTGAGTGGACCTTCTGATACTGCTCCTCCTTGCGTGGAAAAGGGGAAAGAAC
TGCATGCATATTATTCAGCGTCCTATATTCAAAGGATATTCTTGGTGATCTTGGAAGTGT
CCGTATC**ATG**GAATCAATCTCTATGATGGGAAGCCCTAAGAGCCTTAGTGAAACTTGTTT
ACCTAATGGCATAAATGGTATCAAAGATGCAAGGAAGGTCACTGTAGGTGTGATTGGAAG
TGGAGATTTTGCCAAATCCTTGACCATTCGACTTATTAGATGCGGCTATCATGTGGTCAT
AGGAAGTAGAAATCCTAAGTTTGCTTCTGAATTTTTTCCTCATGTGGTAGATGTCACTCA
TCATGAAGATGCTCTCACAAAAACAAATATAATATTTGTTGCTATACACAGAGAACATTA
TACCTCCCTGTGGGACCTGAGACATCTGCTTGTGGGTAAAATCCTGATTGATGTGAGCAA
TAACATGAGGATAAACCAGTACCCAGAATCCAATGCTGAATATTTGGCTTCATTATTCCC
AGATTCTTTGATTGTCAAAGGATTTAATGTTGTCTCAGCTTGGGCACTTCAGTTAGGACC
TAAGGATGCCAGCCGGCAGGTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGT
TATTGAACTTGCCCGCCAGTTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGC
CAGAGAGATTGAAAATTTACCCCTACGACTCTTTACTCTCTGGAGAGGGCCAGTGGTGGT
AGCTATAAGCTTGGCCACATTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCC
ATATGCTAGAAACCAACAGAGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAAC
CTTACCTATAGTTGCCATTACTTTGCTCTCCCTAGTATACCTTGCAGGTCTTCTGGCAGC
TGCTTATCAACTTTATTACGGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTG
GTTACAGTGTAGAAAACAGCTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGC
CTACAGCCTCTGCTTACCGATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTA
TCAGCAGGTTCATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGA
AATGTATATCTCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTC
TATCCCTTCAGTGAGCAATGCTTTAAACTGGAGAGAATTCAGTTTTATTCAGTCTACACT
TGGATATGTCGCTCTGCTCATAAGTACTTTCCATGTTTTAATTTATGGATGGAAACGAGC
TTTTGAGGAAGAGTACTACAGATTTTATACACCACCAAACTTTGTTCTTGCTCTTGTTTT
GCCCTCAATTGTAATTCTGGATCTTTTGCAGCTTTGCAGATACCCAGAC**TGA**GCTGGAAC
TGGAATTTGTCTTCCTATTGACTCTACTTCTTTAAAAGCGGCTGCCCATTACATTCCTCA
GCTGTCCTTGCAGTTAGGTGTACATGTGACTGAGTGTTGGCCAGTGAGATGAAGTCTCCT
CAAAGGAAGGCAGCATGTGTCCTTTTTCATCCCTTCATCTTGCTGCTGGGATTGTGGATA
TAACAGGAGCCCTGGCAGCTGTCTCCAGAGGATCAAAGCCACACCCAAAGAGTAAGGCAG
ATTAGAGACCAGAAAGACCTTGACTACTTCCCTACTTCCACTGCTTTTTCCTGCATTTAA
GCCATTGTAAATCTGGGTGTGTTACATGAAGTGAAAATTAATTCTTTCTGCCCTTCAGTT
CTTTATCCTGATACCATTTAACACTGTCTGAATTAACTAGACTGCAATAATTCTTTCTTT
TGAAAGCTTTTAAAGGATAATGTGCAATTCACATTAAAATTGATTTTCCATTGTCAATTA
GTTATACTCATTTTCCTGCCTTGATCTTTCATTAGATATTTTGTATCTGCTTGGAATATA
TTATCTTCTTTTTAACTGTGTAATTGGTAATTACTAAAACTCTGTAATCTCCAAAATATT
GCTATCAAATTACACACCATGTTTTCTATCATTCTCATAGATCTGCCTTATAAACATTTA
AATAAAAAGTACTATTTAATGATTTAACTTCTGTTTTGAAAAAAAAAAAAAAAAAAAAA
AAA

FIG._1

MESISMMGSPKSLSETCLPNGINGIKDARKVTVGVIGSGDFAKSLTIRLIRCGYHVVIGS
RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLLVGKILIDVSNNM
RINQYPESNAEYLASLFPDSLIVKGFNVVSAWALQLGPKDASRQVYICSNNIQARQQVIE
LARQLNFIPIDLGSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA
RNQQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTKYRRFPPWLETWLQ
CRKQLGLLSFFFAMVHVAYSLCLPMRRSERYLFLNMAYQQVHANIENSWNEEEVWRIEMY
ISFGIMSLGLLSLLAVTSIPSVSNALNWREFSFIQSTLGYVALLISTFHVLIYGWKRAFE
EEYYRFYTPPNFVLALVLPSIVILDLLQLCRYPD

FIG._2

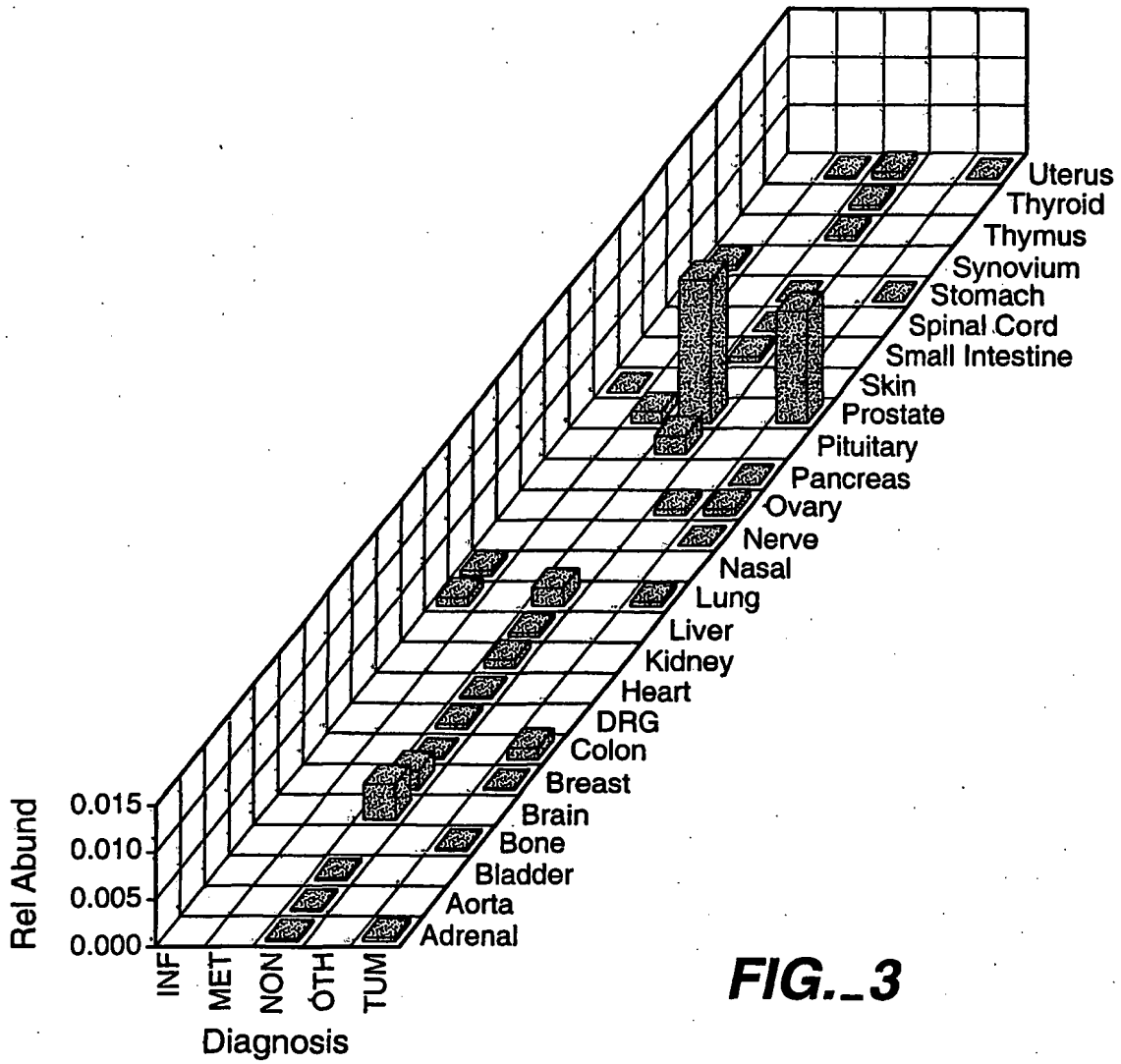
Uterus
Thyroid
Thymus
Synovium
Stomach
Spinal Cord
Small Intestine
Skin
Prostate
Pituitary
Pancreas
Ovary
Nerve
Nasal
Lung
Liver
Kidney
Heart
DRG
Colon
Breast
Brain
Bone
Bladder
Aorta
Adrenal
Rel Abund
0.015
0.010
0.005
0.000
INF
MET
NON
OTH
TUM
Diagnosis

FIG._3

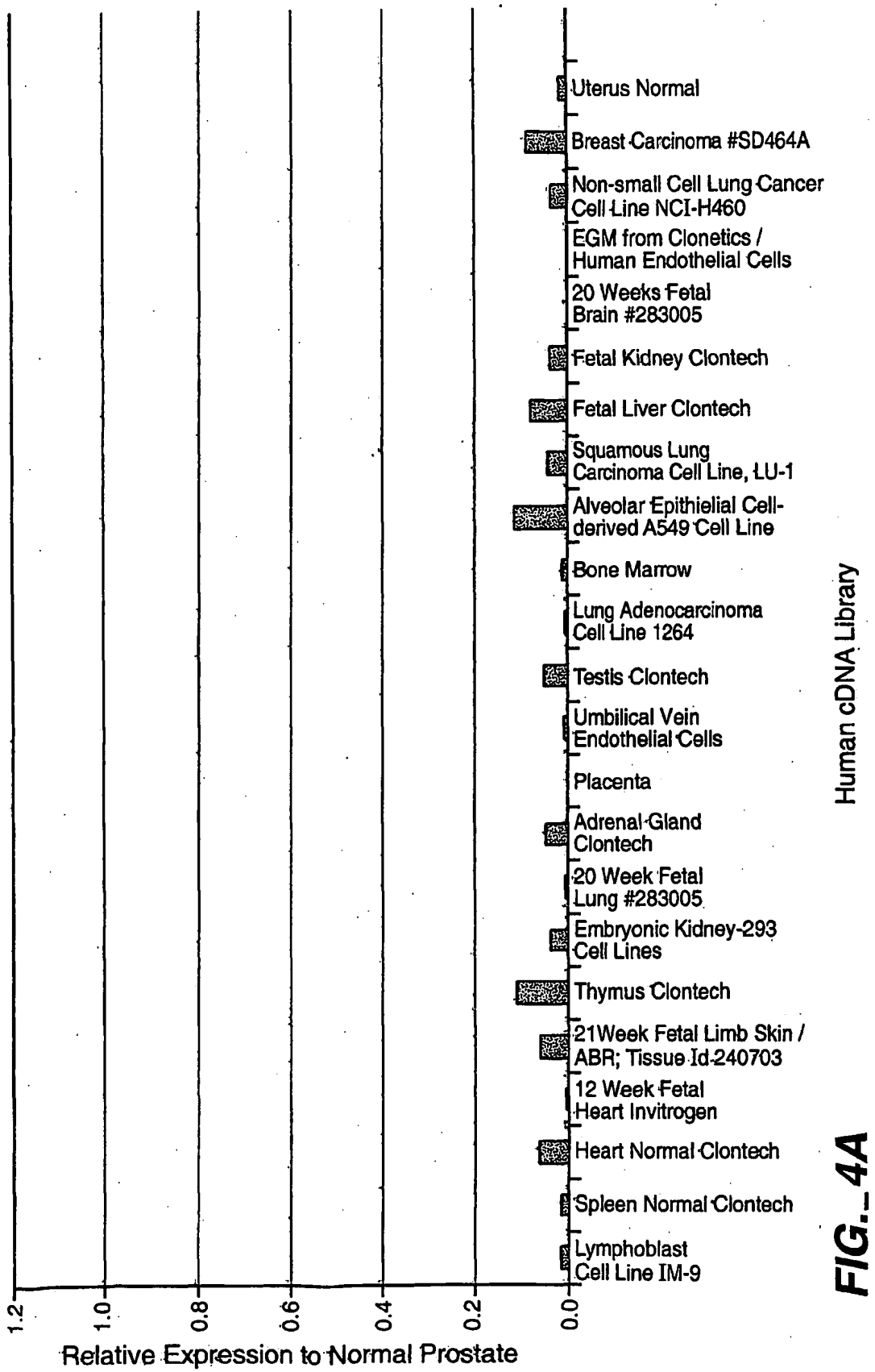
Uterus Normal
Breast Carcinoma #SD464A
Non-small Cell Lung Cancer Cell Line NCI-H460
EGM from Clonetics / Human Endothelial Cells
20 Weeks Fetal Brain #283005
Fetal Kidney Clontech
Fetal Liver Clontech
Squamous Lung Carcinoma Cell Line, LU-1
Alveolar Epithielial Cell-derived A549 Cell Line
Bone Marrow
Lung Adenocarcinoma Cell Line 1264
Testis Clontech
Umbilical Vein Endothelial Cells
Placenta
Adrenal Gland Clontech
20 Week Fetal Lung #283005
Embryonic Kidney-293 Cell Lines
Thymus Clontech
21Week Fetal Limb Skin / ABR; Tissue Id 240703
12 Week Fetal Heart Invitrogen
Heart Normal Clontech
Spleen Normal Clontech
Lymphoblast Cell Line IM-9
Human cDNA Library
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Relative Expression to Normal Prostate

FIG._4A

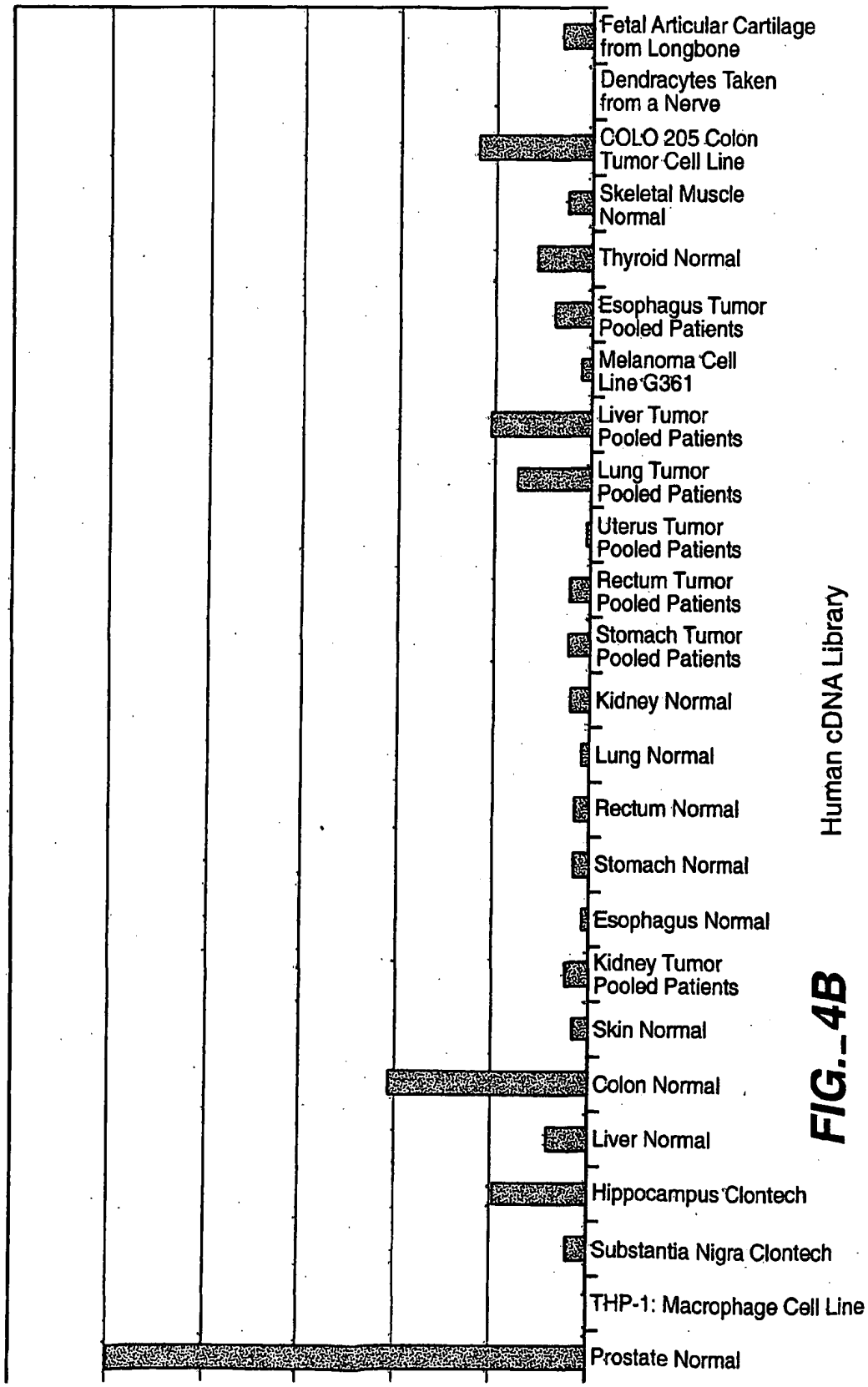
Fetal Articular Cartilage from Longbone
Dendracytes Taken from a Nerve
COLO 205 Colon Tumor Cell Line
Skeletal Muscle Normal
Thyroid Normal
Esophagus Tumor Pooled Patients
Melanoma Cell Line G361
Liver Tumor Pooled Patients
Lung Tumor Pooled Patients
Uterus Tumor Pooled Patients
Rectum Tumor Pooled Patients
Stomach Tumor Pooled Patients
Kidney Normal
Lung Normal
Rectum Normal
Stomach Normal
Esophagus Normal
Kidney Tumor Pooled Patients
Skin Normal
Colon Normal
Liver Normal
Hippocampus Clontech
Substantia Nigra Clontech
THP-1: Macrophage Cell Line
Prostate Normal
Human cDNA Library

FIG._4B

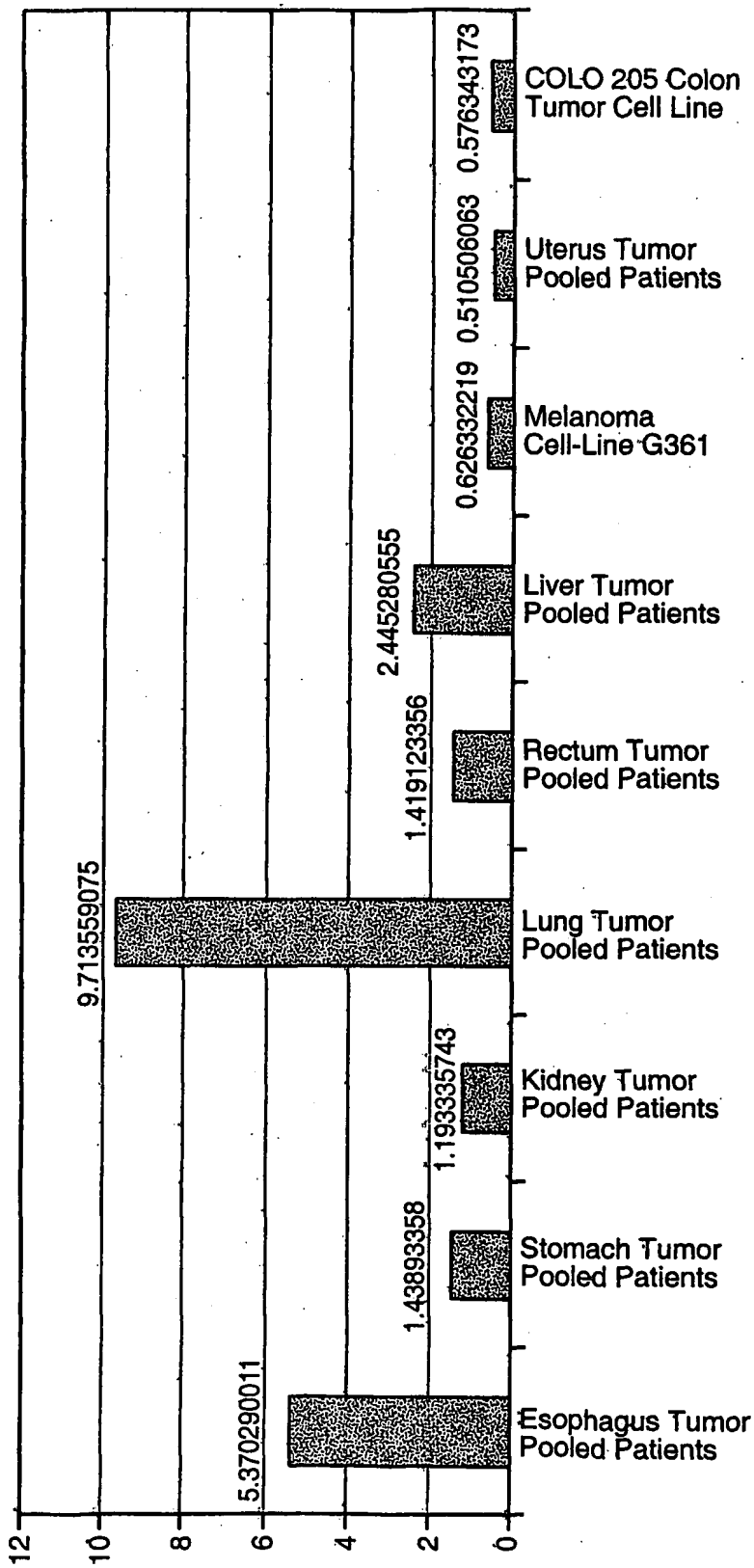
0.576343173
COLO 205 Colon Tumor Cell Line
0.510506063
Uterus Tumor Pooled Patients
0.626332219
Melanoma Cell-Line G361
2.445280555
Liver Tumor Pooled Patients
1.419123356
Rectum Tumor Pooled Patients
9.713559075
Lung Tumor Pooled Patients
1.193335743
Kidney Tumor Pooled Patients
1.43893358
Stomach Tumor Pooled Patients
5.370290011
Esophagus Tumor Pooled Patients
12
10
8
6
4
2
0
cDNA Library

FIG._5

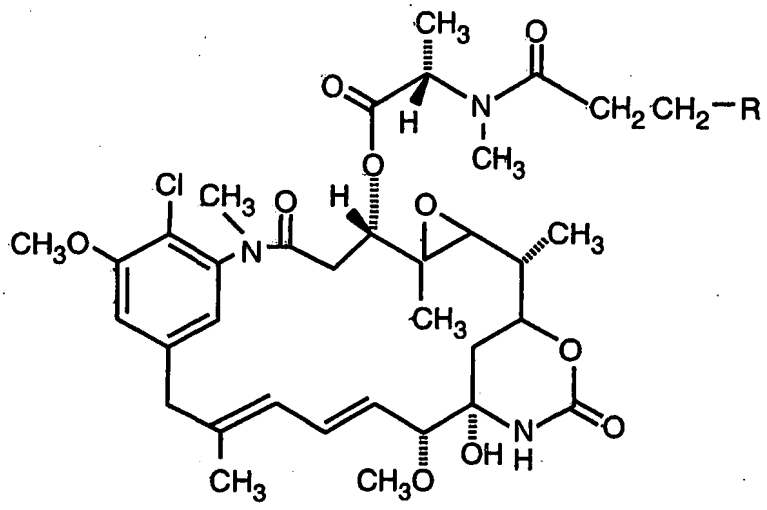
CH3
O
O
N
H
CH2 CH2 R
CH3
O
Cl
CH3
O
H
O
CH3O
N
CH3
CH3
O
N
OH H
O
CH3
CH3O

FIG._6